

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/100934/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Basu, Amitabh and Oertel, Timm ORCID: <https://orcid.org/0000-0001-5720-8978> 2017. Centerpoints: a link between optimization and convex geometry. SIAM Journal on Optimization 27 (2) , pp. 866-889. 10.1137/16M1092908 file

Publishers page: <http://dx.doi.org/10.1137/16M1092908>  
<<http://dx.doi.org/10.1137/16M1092908>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# CENTERPOINTS: A LINK BETWEEN OPTIMIZATION AND CONVEX GEOMETRY\*

AMITABH BASU<sup>†</sup> AND TIMM OERTEL<sup>‡</sup>

**Abstract.** We introduce a concept that generalizes several different notions of a “centerpoint” in the literature. We develop an oracle-based algorithm for convex mixed-integer optimization based on centerpoints. Further, we show that algorithms based on centerpoints are “best possible” in a certain sense. Motivated by this, we establish structural results about this concept and provide efficient algorithms for computing these points. Our main motivation is to understand the complexity of oracle-based convex mixed-integer optimization.

**Key words.** convex mixed-integer optimization, centerpoints, convex and discrete geometry

**AMS subject classifications.** 52A35, 52C07, 90C11, 90C25

**DOI.** 10.1137/16M1092908

**1. Introduction.** Consider the following unconstrained optimization problem

$$(1) \quad \min_{(x,y) \in \mathbb{Z}^n \times \mathbb{R}^d} g(x,y),$$

where  $g : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}$  is a convex function. To keep the problem as general as possible, we assume that  $g$  can be accessed only by a first-order evaluation oracle. In other words, when queried at a point  $(x, y)$ , the oracle returns the corresponding function value  $g(x, y)$  and an element from the subdifferential  $\partial g(x, y)$ . This allows us to model very general, possibly nonsmooth, convex functions. The only additional assumption we make to keep (1) tractable is that the minimization problem is bounded.

We present a general cutting plane algorithm based on the concept of *centerpoints*, which we define below. We call it the *centerpoint algorithm*. Our approach bears similarities to a number of continuous convex minimization algorithms and to Lenstra-type algorithms [31, 20] for convex integer optimization problems. Most variations of Lenstra-type algorithms rely on a combination of the ellipsoid method and enumeration on lower-dimensional subproblems [27, 24, 33, 25, 38, 13]. The key difference is that our algorithm avoids enumerating low-dimensional subproblems.

The main feature of this approach is that, from the point of view of the number of function oracle calls, this algorithm is best possible, up to a lower-order factor. For this purpose, we present our results for a somewhat general convex optimization problem (see section 2 for details), and then specialize the results to continuous/integer/mixed-integer convex optimization. We now proceed to the central concept of this paper.

**Centerpoints.** Let  $\mu$  be a Borel measure on  $\mathbb{R}^n$  such that  $0 < \mu(\mathbb{R}^n) < \infty$ . Without any loss of generality, we normalize the measure to be a probability measure,

---

\*Received by the editors September 8, 2016; accepted for publication (in revised form) March 21, 2017; published electronically May 9, 2017. Parts of the results presented in this paper appear in the second author’s Ph.D. thesis, prepared at ETH Zurich, Zurich [37] and in the Proceedings of the International Conference on Integer Programming and Combinatorial Optimization, Springer, Cham, Switzerland, 2016 [4].

<http://www.siam.org/journals/siopt/27-2/M109290.html>

**Funding:** Amitabh Basu was supported in part by NSF grant CMMI1452820.

<sup>†</sup>Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD 21218 (basu.amitabh@jhu.edu).

<sup>‡</sup>School of Mathematics, Cardiff University, Cardiff, Wales, UK (oertelt@cardiff.ac.uk).

i.e.,  $\mu(\mathbb{R}^n) = 1$ . A reader unfamiliar with measure theory may simply consider  $\mu$  to be a mixed-integer version of the Lebesgue measure, which generalizes both the standard volume and the “counting measure” for the integer lattice. See (7) for a precise definition.

For  $S \subseteq \mathbb{R}^n$  nonempty and closed, we define the set of centerpoints  $\mathcal{C}(S, \mu) \subseteq S$  as the set that attains the following maximum:

$$(2) \quad \mathcal{F}(S, \mu) := \max_{x \in S} \inf_{u \in \mathcal{S}^{n-1}} \mu(H^+(u, x)),$$

where  $\mathcal{S}^{n-1}$  denotes the  $(n-1)$ -dimensional unit sphere with center in the origin and  $H^+(u, x)$  denotes the half-space  $\{y \in \mathbb{R}^n : u \cdot (y - x) \geq 0\}$ . In other words,  $\mathcal{F}(S, \mu)$  is the largest real number  $M > 0$ , such that there is a point  $x \in S$  with the property that any half-space containing  $x$  has measure at least  $M$ . This definition unifies several definitions from convex geometry, computer science, and statistics. Two notable examples are the following:

1. *Winternitz measure of symmetry.* Let  $\mu$  be the Lebesgue measure restricted to a convex body  $K$  (i.e.,  $K$  is compact and has a nonempty interior) or, equivalently, the uniform probability measure on  $K$ , and let  $S = \mathbb{R}^n$ .  $\mathcal{F}(S, \mu)$  in this setting is known in the literature as the *Winternitz measure of symmetry* of  $K$ , and the centerpoints  $\mathcal{C}(S, \mu)$  are the “points of symmetry” of  $K$ . This notion was studied by Grünbaum in [22] and surveyed by the same author in [23]. Caplin and Nalebuff [10] generalize Grünbaum’s results to measures  $\mu$  with a concave density supported on a compact set  $K$ . Reference [42] is a recent survey on measures of symmetry of convex bodies. Convex geometry literature also studies the closely related concepts of floating bodies and illumination bodies; see [46] for a survey.
2. *Tukey depth and median.* In statistics and computational geometry, the function  $f_\mu : \mathbb{R}^n \rightarrow \mathbb{R}$  defined as

$$(3) \quad f_\mu(x) := \inf_{u \in \mathcal{S}^{n-1}} \mu(H^+(u, x))$$

is known as the *half-space depth function* or the *Tukey depth function* for the measure  $\mu$ , first introduced by Tukey [43]. Taking  $S = \mathbb{R}^n$ , the centerpoints  $\mathcal{C}(\mathbb{R}^n, \mu)$  are known as the *Tukey medians* of the probability measure  $\mu$ , and  $\mathcal{F}(\mathbb{R}^n, \mu)$  is known as the maximum *Tukey depth* of  $\mu$ . Tukey introduced the concept when  $\mu$  is a finite sum of Dirac measures (i.e., a finite set of points with the counting measure), but the concept has been generalized to other probability measures and analyzed from structural [17, 40, 29], as well as computational perspectives [40, 11, 9, 18]. See [34] for a survey of structural aspects and other notions of “depth” used in statistics, and [18] and the references therein for a survey and recent approaches to computational aspects of the Tukey depth.

**Our results.** To the best of our knowledge, all related notions of centerpoints in the literature have only considered the case where the set  $S$  is  $\mathbb{R}^n$ , i.e., the centerpoint can be any point from the Euclidean space. We consider a more general  $S$ , as this allows us to analyze convex optimization problems where the solutions have to satisfy side constraints like mixed-integer constraints, sparsity constraints (e.g., compressed sensing), or complementarity constraints. Essentially, the closed subset  $S$  is going to represent nonconvex feasibility constraints of our optimization problem; at the most general level, all we require from  $S$  is that it is closed and nonempty.

TABLE 1  
Best bounds for the convex optimization problem (4) with box constraints.

$S$	Upper bound	Lower bound	
$\mathbb{R}^d$	$O\left(\log_{\frac{e}{e-1}}\left(\frac{B^d}{\delta}\right)\right)$	$\Omega\left(\log_2\left(\frac{B^d}{\delta}\right)\right)$	(Nemirovsky and Yudin)
$\mathbb{Z}^n$	$O(n2^n \log_2(B))$	$\Omega(2^n \log_2(B))$	
$\mathbb{Z}^n \times \mathbb{R}^d$	$O\left(2^n(d+1) \log_2\left(\frac{B^{n+d}}{\delta}\right)\right)$	$\Omega\left(2^n \left(\log_2 \frac{B^d}{\delta}\right)\right)$	

In section 2, we elaborate on this connection between centerpoints and algorithms for optimizing convex functions over general closed sets  $S$ . We first give an algorithm for solving such problems given access to first-order (separation) oracles, based on centerpoints. We then focus on convex *mixed-integer* optimization and show that the centerpoint algorithms is “best possible” in a certain sense, amongst a large class of first-order oracle-based methods—see Table 1. This comprises our main motivation in studying centerpoints.

In section 3, we provide lower bounds on the value of  $\mathcal{F}(S, \mu)$ . In section 3.1, we obtain lower bounds in terms of the Helly number of  $S$  with minimal assumptions on  $S$  and  $\mu$ . In section 3.2, we obtain better lower bounds for the special case when  $S = \mathbb{Z}^n \times \mathbb{R}^d$  and  $\mu$  is the “mixed-integer” uniform measure on  $K \cap (\mathbb{Z}^n \times \mathbb{R}^d)$ , where  $K$  is some convex body. Such bounds immediately imply bounds on the complexity of oracle-based convex mixed-integer optimization algorithms.

In section 4, we present a number of exact and approximation algorithms for computing centerpoints. To the best of our knowledge, the computational study of centerpoints has only been done for measures  $\mu$  that are a finite sum of Dirac measures, i.e., for finite point sets, or when  $\mu$  is the uniform measure on two-dimensional polygons (e.g., see [8] and the references therein). We initiate a study for other measures; in particular, the uniform measure on a convex body, the counting measure on the integer points in a convex body, and the mixed-integer volume of the mixed-integer points in a convex body. All our algorithms are exponential in the dimension  $n$  but polynomial in the remaining input data, so these are polynomial time algorithms if  $n$  is assumed to be a constant. Algorithms that are polynomial in  $n$  are likely to not exist because of the reduction to the so-called *closed hemisphere problem*—see Chapter 7 by Bremner, Fukuda, and Rosta in the collection of articles in [32].

We mention that the algorithms for computing centerpoints from section 4 are based on the standard Turing machine model of computation and, therefore, work with rational arithmetic. Consequently, since the coordinates of a centerpoint could be irrational, our algorithms return points whose coordinates are arbitrary close approximations of the centerpoint coordinates. We alert the reader that our analysis of the oracle complexity of cutting plane algorithms in section 2 ignores such arithmetic issues and the results assume that exact centerpoints are used in the *optimization algorithm*. A general framework for handling such arithmetic issues is described in [20]. For this reason, we do not discuss them further in this manuscript.

Throughout this paper we use the notation  $H^-(u, x) := \{y \in \mathbb{R}^n : u \cdot (y - x) \leq 0\}$  and  $H(u, x) := \{y \in \mathbb{R}^n : u \cdot (y - x) = 0\}$  for  $u \in \mathcal{S}^{n-1}$  and  $x \in \mathbb{R}^n$ . Recall that  $H^+(u, x)$  denotes the half-space  $\{y \in \mathbb{R}^n : u \cdot (y - x) \geq 0\}$ . With  $\text{int}(X)$  we will denote the interior of  $X \subseteq \mathbb{R}^n$ .

**2. The connection to optimization.** Given a nonempty, closed set  $S \subseteq \mathbb{R}^n$ , consider the following optimization problem:

$$(4) \quad \min_{x \in S} g(x),$$

where  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function given by a first-order evaluation oracle. The same setting is currently being studied by De Loera et al., calling it *S-optimization* [14]. Related models have also been proposed in integer programming; see [12] and [47] and the references therein. We first define the class of algorithms against which we will compare the centerpoint algorithm. We refer to this class as *cutting plane algorithms*.

**DEFINITION 2.1** (cutting plane algorithm). *Let  $S \subseteq \mathbb{R}^n$  be a nonempty, closed set and let  $\nu$  be a Borel measure on  $\mathbb{R}^n$  such that  $\nu(\mathbb{R}^n \setminus S) = 0$ , i.e.,  $\nu$  is supported on  $S$ . A cutting plane algorithm for (4) with stopping criterion based on  $\nu$  is an algorithm with the following structure:*

**INPUT:** *An error guarantee  $\delta > 0$  and a convex set  $E_0$  that contains the optimal solution in its interior and with  $\nu(E_0) > 0$ .*

**ITERATIONS:** *At each iteration  $i = 1, 2, \dots$  the algorithm selects a point  $x_i \in \text{int}(E_{i-1}) \cap S$ , and then makes a call to a first-order oracle for  $g$  at  $x_i$ , which returns the function value  $g(x_i)$  and a subgradient  $h_i \in \partial g(x_i)$ . We define  $x^* := \arg\min_{x \in \{x_1, \dots, x_i\}} g(x)$  and define  $E_i$  such that*

$$(5) \quad E_i \supseteq \{x \in E_0 : g(x^*) - g(x_j) \geq h_j \cdot (x - x_j) \ \forall j = 1, \dots, i\}.$$

**OUTPUT:** *The cutting plane algorithm stops at the  $N$ th iteration when  $\nu(\text{int}(E_N)) \leq \delta$  and returns  $x^*$ .*

When  $S = \mathbb{R}^n$  and  $\nu$  is the standard Lebesgue measure, we obtain the standard cutting plane algorithms for continuous convex optimization, such as the ellipsoid method, method of centers of gravity, the Kelly method, or the level method [36, section 3]. A variant of the cutting plane method which utilizes random sampling was explored by Bertsimas and Vempala in [7] for continuous convex optimization. Their method also falls under the general framework of Definition 2.1. When  $S = \mathbb{Z}^n$  and  $\nu$  is the counting measure for  $\mathbb{Z}^n$ , we obtain cutting plane algorithms for convex integer optimization problems. When  $S = \mathbb{Z}^n \times \mathbb{R}^d$ , we obtain algorithms for convex mixed-integer optimization. A natural choice of the measure is the mixed-integer measure  $\bar{\mu}_{\text{mixed}}$ , which we will discuss in the next section.

*Remark 2.2.* We give a short justification of our stopping criterion in Definition 2.1. Given general convex functions and the absence of any known structure for  $S$ , one can only guarantee an approximate solution from an algorithmic point of view in general (for structured  $S$  like  $\mathbb{Z}^n$ , the situation is different). Typically, the quality of such approximations is then quantified by an additive or multiplicative gap with respect to the optimal function value. These often require additional estimations based on further parameters, for example, a Lipschitz constant. Instead of considering the gap in the function values, we approximate an optimal point and we quantify the approximation quality by an appropriate measure  $\nu$ , thus circumventing any additional assumptions or sources of error such as Lipschitz constants. At the end of this section, we elaborate on how to extend our results to derive bounds on the additive gap with respect to the optimal function value in the mixed-integer case.

*Remark 2.3.* We also briefly comment on the assumption that  $\nu(\mathbb{R}^n \setminus S) = 0$ , i.e., the measure is supported on  $S$ . One could consider more general measures that do not satisfy this condition and analyze the class of algorithms obtained thus. However, the mathematical analysis of this more general situation becomes more tedious with many corner cases that need to be handled, without adding any new insight. It is more elegant to restrict the analysis to measures that are supported on  $S$ . Consequently, we build this into the definition of our cutting plane algorithm.

**DEFINITION 2.4** (centerpoint algorithm). *Let  $S \subseteq \mathbb{R}^n$  be a nonempty, closed set and let  $\nu$  be a Borel measure on  $\mathbb{R}^n$  such that  $\nu(\mathbb{R}^n \setminus S) = 0$ . The centerpoint algorithm is a cutting plane algorithm for (4) that chooses  $x_i \in \mathcal{C}(S, \nu_i)$  from the set of centerpoints, where  $\nu_i$  is the measure  $\nu$  restricted to  $\text{int}(E_{i-1})$ , and defines  $E_i$  to be the right-hand side of (5).*

For our general bounds we need three parameters related to  $S$  and  $\nu$ .

**DEFINITION 2.5.** *Let  $S \subseteq \mathbb{R}^n$  be a nonempty, closed subset and  $\nu$  be a measure on  $\mathbb{R}^n$  that is finite for any bounded set.*

- (i) *For any bounded convex set  $C \subseteq \mathbb{R}^n$  with  $\nu(C) > 0$ , define  $\nu_C$  as the normalized finite measure  $\nu$  restricted to  $C$ . We define*

$$c(S, \nu) := \inf_{\nu_C} \mathcal{F}(S, \nu_C).$$

- (ii) *We define the degeneracy parameter*

$$\chi(S, \nu) := \max_{y \in S} \min_{u \in S^{n-1}} \nu(\{x \in \mathbb{R}^n : u \cdot x = u \cdot y\}).$$

- (iii) *We say that  $\nu$  is lower semicontinuous if for every  $\epsilon > 0$  and every open set  $A \subseteq \mathbb{R}^n$ , there exists a closed set  $A' \subseteq A$  such that*

$$\nu(\text{int}(A')) \geq \nu(A) - \epsilon.$$

In Theorem 3.3, we will show that  $c(S, \nu) \geq h(S)^{-1}$ , where  $h(S)$  denotes the Helly number of  $S$ . However, for certain types of measures one can obtain stronger bounds, e.g., see Corollary 2.8. Note that the parameter  $\chi$  is zero when  $S$  corresponds to the whole space, a mixed-integer lattice, or to sets obtained from sparsity constraints on the variables. However, if  $S = \mathbb{Z}^n$  and  $\nu$  corresponds to the counting measure, then  $\chi(S, \nu) = 1$ .

Our first general result showing the asymptotic optimality of the centerpoint-based algorithm amongst cutting plane algorithms for (4) is the following.

**THEOREM 2.6** (general optimality bounds). *Let  $S$  be a nonempty, closed set. Let  $\nu$  be a measure such that  $\nu(\mathbb{R}^n \setminus S) = 0$ ,  $\nu$  is finite for bounded sets, and  $\nu$  is lower semicontinuous (as defined in part (iii) of Definition 2.5). Further, assume that  $c(S, \nu) > 0$ . Let  $\delta > 0$  and  $E_0 \subseteq \mathbb{R}^n$  with  $\nu(E_0) = V > 0$ . The centerpoint algorithm makes at most*

$$\left\lceil \log_{\frac{1}{1-c(S, \nu)}} \left( \frac{V}{\delta} \right) \right\rceil$$

*first-order oracle calls for any convex function  $g$ . Moreover, for any cutting-plane-based algorithm  $\mathcal{A}$  with a stopping criterion based on  $\nu$ , there exists a convex function  $\hat{g}$  such that  $\mathcal{A}$  will make at least*

$$\left\lceil \left( \log_2 \left( \frac{V}{\delta + \chi(S, \nu)} \right) \right) \right\rceil - 1$$

*first-order evaluation oracle calls to  $\hat{g}$ .*

*Proof.* The upper bound follows from the fact that by choosing the centerpoint at every iteration, one can guarantee that  $\nu(\text{int}(E_i)) \leq (1 - c(S, \nu))\nu(\text{int}(E_{i-1}))$  for every  $i = 1, \dots, N$ , where  $N$  is the number of iterations such that  $\nu(\text{int}(E_N)) \leq \delta$ .

For the lower bound, it suffices to establish the following claim whose proof appears below.

**CLAIM 2.7.** *For any cutting plane algorithm  $\mathcal{A}$  and a real number  $\epsilon > 0$ , there exists a sequence of convex functions  $\{g_i\}_{i=1}^\infty$  such that for every  $i \geq 1$ , if  $\mathcal{A}$  runs for  $i$  iterations on  $g_i$ , then  $\nu(\text{int}(E_i)) \geq (\frac{1}{2})^i \nu(\text{int}(E_0)) - 2(\epsilon - \chi(S, \mu))$ .*

To obtain the lower bound stated in the theorem from this, we do the following: Given  $\delta > 0$ ,  $V > 0$ , set  $N^* := \lceil (\log_2(\frac{V}{\delta + \chi(S, \nu)})) \rceil - 1$  and we run  $\mathcal{A}$  on  $g_{N^*}$ . Using Claim 2.7 with  $\epsilon = \frac{\delta}{2}$ , we obtain that  $\nu(\text{int}(E_{N^*})) \geq (\frac{1}{2})^{N^*} \nu(\text{int}(E_0)) - \delta - 2\chi(S, \mu)$ . Since the stopping criterion for the algorithm is  $\nu(\text{int}(E_{N^*})) \leq \delta$ , the inequality implies that  $\mathcal{A}$  requires at least  $N^*$  iterations to stop.

*Proof of Claim 2.7.* We construct the sequence  $\{g_i\}_{i=1}^\infty$  in an adversarial manner; we will actually construct  $\text{epi}(g_i)$ , the epigraphs of  $g_i$ ,<sup>1</sup> and use the fact that an epigraph defines a convex function uniquely and vice versa.

In fact, we will inductively construct three sequences: convex functions  $\{g_i\}_{i=0}^\infty$ , vectors  $\{h_i\}_{i=1}^\infty \subseteq \mathbb{R}^n$ , and real numbers  $\{\xi_i\}_{i=1}^\infty \subseteq \mathbb{R}$  such that the following conditions hold for every  $i \geq 1$ .

- (i) If algorithm  $\mathcal{A}$  runs on the function  $g_i$  for  $i$  iterations,

$$\nu(\text{int}(E_i)) \geq \left(\frac{1}{2}\right)^i \nu(\text{int}(E_0)) - \left(\sum_{j=0}^{i-1} \frac{1}{2^j}\right) \cdot (\epsilon + \chi(S, \nu)),$$

- (ii) For any  $0 \leq j \leq i$ , when the algorithm is executed for  $j$  iterations on  $g_j$  and  $g_i$ , it queries the same points  $x_1, \dots, x_j$ .
- (iii) For any  $0 \leq j \leq i$ ,  $\{h_j\} = \partial g_i(x_j)$ ,
- (iv) Let  $x_1, \dots, x_i$  be the points queried by  $\mathcal{A}$  when executed on  $g_i$  for  $i$  iterations.

Then,

$$\text{epi}(g_i) = \bigcap_{j=1}^i \left\{ (x, t) : E_0 \times \mathbb{R} : t \geq h_j \cdot (x - x_j) - \sum_{k=1}^j \xi_k \right\}.$$

First observe that (i) of the claim implies  $\nu(\text{int}(E_i)) \geq (\frac{1}{2})^i \nu(\text{int}(E_0)) - (\sum_{j=0}^{i-1} \frac{1}{2^j}) \cdot (\epsilon + \chi(S, \nu)) \geq (\frac{1}{2})^i \nu(\text{int}(E_0)) - 2(\epsilon + \chi(S, \mu))$ . This would complete the proof of Claim 2.7.

We prove the claim inductively. Let  $x_1 \in \text{int}(E_0)$  be the first point queried by  $\mathcal{A}$  on any convex function. Choose  $h_1 \in \mathbb{R}^n \setminus \{0\}$  such that  $\nu(\{x \in \mathbb{R}^n : h_1 \cdot x = h_1 \cdot x_1\}) \leq \chi(S, \nu)$ , and that  $\nu(\text{int}(E_0) \cap \{x : h_1 \cdot (x - x_1) \leq 0\}) \geq \frac{1}{2} \nu(\text{int}(E_0))$ . Such a choice of  $h_1$  always exists. Set  $\xi_1 = 0$ . Finally, define  $\text{epi}(g_1) := \{(x, t) : E_0 \times \mathbb{R} : t \geq h_1 \cdot (x - x_1)\}$ . One can now check that (i)–(iv) in the above claim hold for  $i = 1$ .

Now, suppose we have defined  $g_1, \dots, g_i$ ,  $h_1, \dots, h_i$ , and  $\xi_1, \dots, \xi_i$  for some  $i \geq 1$  such that (i)–(iv) hold. We now construct  $g_{i+1}$ ,  $h_{i+1}$ , and  $\xi_{i+1}$ . Note that by (iii), the algorithm chooses  $x_{i+1} \in \text{int}(E_i)$ , where

$$E_i \supseteq \hat{E}_i := \left\{ x \in E_0 : h_j \cdot (x - x_j) \leq - \sum_{k=j+1}^i \xi_k, j = 1, \dots, i \right\}.$$

<sup>1</sup>For any function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , the epigraph of  $f$  is  $\text{epi}(f) := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : t \geq f(x)\}$ .

If  $x_{i+1} \notin \text{int}(\widehat{E}_i)$ , then set  $g_{i+1} = g_i$ ,  $h_{i+1} \in \partial g_i(x_{i+1})$  (which is well-defined) and  $\xi_{i+1} = 0$ . Otherwise, by our assumption that  $\nu$  is lower semicontinuous, one can choose  $\xi_{i+1} > 0$ , such that following conditions hold for  $\tilde{E}_i := \{x \in E_0 : h_j \cdot (x - x_j) \leq -\sum_{k=j+1}^i \xi_k - \xi_{i+1}, j = 1, \dots, i\}$ :

- (A)  $\nu(\text{int}(\widehat{E}_i)) - \epsilon \leq \nu(\text{int}(\tilde{E}_i))$  and
- (B)  $x_i \in \tilde{E}_i$ .

Let  $h_{i+1} \in \mathbb{R}^n \setminus \{0\}$  such that the following all hold:

- (a)  $\nu(\{x \in \mathbb{R}^n : h_{i+1} \cdot x = h_{i+1} \cdot x_{i+1}\}) \leq \chi(S, \nu)$ ,
- (b)  $\nu(\text{int}(\widehat{E}_i) \cap \{x \in \mathbb{R}^n : h_{i+1} \cdot (x - x_{i+1}) \leq 0\}) \geq \frac{1}{2}\nu(\text{int}(\widehat{E}_i))$  and
- (c)  $\text{int}(\{(x, t) : t \geq h_{i+1} \cdot (x - x_{i+1}) - \sum_{k=1}^{i+1} \xi_k\}) \supseteq \widehat{E}_i \times \{\sum_{k=1}^i \xi_k\}$ .

Condition (a) can be ensured by the definition of  $\chi(S, \nu)$ , (b) can be ensured by choosing one of the closed half-spaces corresponding to the normal minimizing the degeneracy parameter  $\chi(S, \nu)$ , and (c) can finally be ensured by scaling down  $h_{i+1}$  as required. Last, define

$$(6) \quad \text{epi}(g_{i+1}) = \text{epi}(g_i) \cap \left\{ (x, t) : E_0 \times \mathbb{R} : t \geq h_{i+1} \cdot (x - x_{i+1}) - \sum_{k=1}^{i+1} \xi_k \right\}.$$

To confirm condition (i), we observe that  $E_{i+1} \supseteq \widehat{E}_{i+1}$  and, therefore,

$$\begin{aligned} \nu(\text{int}(E_{i+1})) &\geq \nu(\text{int}(\widehat{E}_{i+1})) \\ &= \nu(\text{int}(\tilde{E}_i \cap H_{i+1})) \\ &\geq \nu(\text{int}(\widehat{E}_i \cap H_{i+1})) - \epsilon \\ &\geq \nu(\text{int}(\tilde{E}_i) \cap H_{i+1}) - \chi(S, \nu) - \epsilon \\ &\geq \frac{1}{2}\nu(\text{int}(\widehat{E}_i)) - \chi(S, \nu) - \epsilon. \end{aligned}$$

To verify condition (ii), by induction we simply need to verify that if  $\mathcal{A}$  queries  $x_1, \dots, x_i$  on the first  $i$  iterations while executing on  $g_i$ , then  $g_i(x_k) = g_{i+1}(x_k)$  for all  $k = 1, \dots, i$ . This follows from condition (c) above that was maintained during the choice of  $h_{i+1}$ . Condition (iii) also follows from condition (c) above that was maintained during the choice of  $h_{i+1}$ . Condition (iv) follows from (6) and the fact that inductively condition (ii) ensures that  $\mathcal{A}$  queries the same points  $x_1, \dots, x_i$  on  $g_i$  and  $g_{i+1}$ .  $\square$

This concludes the proof of Theorem 2.6.  $\square$

We obtain, as a special case, a known result of Nemirovski and Yudin [35] on the optimality of the centerpoint algorithm for continuous convex optimization.

**COROLLARY 2.8** (continuous convex optimization bounds). *The centerpoint algorithm is optimal amongst cutting plane algorithms for continuous convex optimization in terms of number of function oracle calls, up to the constant factor  $\log_2(\frac{e}{e-1})$ .*

*Proof.* Grunbaüm showed that when  $S = \mathbb{R}^n$  and  $\nu$  is the Lebesgue measure,  $c(S, \nu) = (\frac{n}{n+1})^n \geq \frac{1}{e}$  [22]. Theorem 2.6 then gives the result.  $\square$

**Remark 2.9.** The assumption  $c(S, \nu) > 0$  in Theorem 2.6 excludes pathological situations where the algorithm does not terminate. As an example, if  $S = \mathcal{S}^{n-1}$  is the unit sphere,  $\nu$  is the uniform measure supported on  $S$ , and  $g = \|x\|_2$ , then at every iteration of the algorithm, only a single point is excluded and no progress is made in terms of the measure even after countably many iterations.



**A more refined analysis for lattices and mixed-integer lattices.** In this section we consider the two cases  $S = \mathbb{Z}^n$  and  $S = \mathbb{Z}^n \times \mathbb{R}^d$  and where  $\mu$  is the “uniform measure” on a convex set intersected with  $S$ . More precisely, let  $K \subseteq \mathbb{R}^n \times \mathbb{R}^d$  be a convex set. Let  $\text{vol}_d$  be the  $d$ -dimensional volume (Lebesgue measure). We define the *mixed-integer volume with respect to  $K$*  as

$$(7) \quad \mu_{\text{mixed},K}(C) := \frac{\sum_{z \in \mathbb{Z}^n} \text{vol}_d(C \cap K \cap (\{z\} \times \mathbb{R}^d))}{\sum_{z \in \mathbb{Z}^n} \text{vol}_d(K \cap (\{z\} \times \mathbb{R}^d))}$$

for any Lebesgue measurable subset  $C \subseteq \mathbb{R}^n \times \mathbb{R}^d$ . For later use we want to introduce the notation  $\bar{\mu}_{\text{mixed}}(C) = \sum_{z \in \mathbb{Z}^n} \text{vol}_d(C \cap (\{z\} \times \mathbb{R}^d))$ . The dimensions  $n$  and  $d$  will be clear from the context.

*Remark 2.10.* Let  $K \subseteq \mathbb{R}^{n+d}$  be a convex body and let  $\mu_{\text{mixed},K}$  denote the mixed-integer volume with respect to  $K$ , as defined in (7). Observe that, if  $n = 0$ , then  $\mu_{\text{mixed},K}(H^+(u, x))$  is continuous in  $u$ . Thus, the infimum over the compact unit sphere is achieved. When  $n \neq 0$  the function  $\mu_{\text{mixed},K}(H^+(u, x))$  remains continuous nearly everywhere on  $\mathcal{S}^{n+d-1}$ . Only on  $\mathcal{S}^{n-1} \times \{0\}^d$  is the function piecewise constant. In particular, this implies that the infimum in (2) and (3) is actually achieved.

We will show below that when  $S = \mathbb{Z}^n$ , and  $\nu$  is the counting measure on  $\mathbb{Z}^n$ ,  $c(S, \nu) = \frac{1}{2^n}$ ; see Corollary 3.4. Note that when  $S = \mathbb{Z}^n$ , one can choose  $\delta < 1$  in which case a cutting plane algorithm will return a true optimal solution because if  $\nu(\text{int}(E_N)) \leq \delta$ , this means there is no integer point left in  $E_N$  and thus  $x^*$  must be an optimal solution. To this point we have made no assumption on our initial  $E_0$ , except that  $\nu(E_0) = V > 0$ . It is possible to design  $E_0$  such that either the lower or the upper bound provided by Theorem 2.6 is the best possible. Examples would be  $E_0 = [0, B] \times \{0\}^{n-1}$  or  $E_0 = [0, B] \times [0, 1]^{n-1}$ , respectively, where  $B \geq 1$ . However, these are rather artificial constructions. A more common assumption is that an optimal solution  $\hat{x}$  has a bounded representation, say  $\|\hat{x}\|_\infty \leq B$  for some natural number  $B \geq 1$ . This would imply that we initiate the algorithm with  $E_0$  being a box. For ease of presentation we will assume that  $E_0 = [0, B]^n$ , i.e.,  $E_0$  is not centrally symmetric to the origin. It follows, that  $\nu(E_0)$  equals  $B^n$  and  $B^{n+d}$  for the integer and the mixed-integer case, respectively. (Of course other definitions of initial  $E_0$ 's are also plausible, for example, balls. In these cases one could also do a more refined analysis as described below. In the case of the ball the bound would differ only by a linear factor in terms of the root of the dimension. This is a consequence of John's ellipsoid theorem.) Then, the bound in Theorem 2.6 says that the centerpoint-based algorithm takes at most

$$\left\lceil \log_{\frac{1}{1-c(S, \nu)}} \left( \frac{V}{\delta} \right) \right\rceil = O(n 2^n \log_2(B))$$

function oracle calls, where one uses the inequality  $-\ln(1-x) \geq x$  for  $0 < x < 1$  to deduce that  $-\ln(1 - \frac{1}{2^n}) \geq \frac{1}{2^n}$ . On the other hand, the lower bound in Theorem 2.6 gives

$$\left\lceil \log_2 \left( \frac{V}{\delta} \right) \right\rceil - 1 = \lceil n \log_2(B) \rceil - 1.$$

This exponential gap between the upper and the lower bounds can be improved using the lattice structure of  $S$ .

**THEOREM 2.11** (pure integer convex optimization bound). *Let  $S = \mathbb{Z}^n$ ,  $\nu$  is the counting measure on  $\mathbb{Z}^n$ ,  $E_0 = [0, B]^n$  where  $B \geq 2$  is an integer, and  $\delta < 1$ . Then*

for any cutting plane algorithm  $\mathcal{A}$  there exists an instance such that  $\mathcal{A}$  makes at least  $2^{n-1}(\lfloor \log_2(B) \rfloor + 1)$  first-order evaluation oracle calls on  $\hat{g}$ .

*Proof.* The proof follows the same idea as in the proof of Theorem 2.6, except that this time we exploit the discrete structure of  $S$ . The important thing to illustrate is the choice of the subgradients  $\{h_i\}_{i=1}^\infty$  from that proof.

For each  $v \in \{0, 1\}^{n-1}$  we define the fiber  $F_v := \{0, 1, \dots, B-1\} \times \{v\}$  and let  $S_0 := \cup_{v \in \{0, 1\}^{n-1}} F_v$ . We now construct the adversarial  $\hat{g}$  in an analogous manner to the proof of Theorem 2.6 by defining the adversarial subgradient half-spaces, or cuts. Whenever an algorithm queries the function oracle on a point  $x \notin S_0$ , then we can always choose the subgradient at  $x$  such that the half-space contains  $S_0$ . Otherwise, if  $x \in S_0$ , we know by definition that  $x \in F_{\bar{v}}$  for some  $\bar{v} \in \{0, 1\}^{n-1}$ . We now choose the subgradient half-space that removes at most half of the remaining points in  $F_{\bar{v}}$  and keeps the remaining fibers  $F_v$ ,  $v \neq \bar{v}$  intact. It then follows that on each of the  $2^{n-1}$  fibers, the algorithm has to perform at least  $\lfloor \log_2(B) \rfloor + 1$  function oracle calls. Therefore, in all, the algorithm must perform at least  $2^{n-1}(\lfloor \log_2(B) \rfloor + 1)$  function oracle calls.  $\square$

For the mixed-integer case  $S = \mathbb{Z}^n \times \mathbb{R}^d$  with the measure  $\nu = \bar{\mu}_{mixed}$ , we will show below that  $c(S, \nu) \geq \frac{1}{2^n(d+1)}$ ; see Corollary 3.4. Similarly to the pure integer case above, assuming we start with  $E_0$  as the box  $[0, B]^n$  for some natural number  $B \geq 2$ , the bound in Theorem 2.6 says that the centerpoint-based algorithm takes at most

$$\left\lceil \log_{\frac{1}{1-c(S, \nu)}} \left( \frac{V}{\delta} \right) \right\rceil = O \left( 2^n(d+1) \log_2 \left( \frac{B^{n+d}}{\delta} \right) \right)$$

function oracle calls. On the other hand, the lower bound in Theorem 2.6 gives

$$\left\lceil \log_2 \left( \frac{V}{\delta} \right) \right\rceil = (n+d) \log_2 \left( \frac{B}{\delta} \right).$$

However, similarly to Theorem 2.11, one can improve the lower bound in the mixed-integer case too.

**THEOREM 2.12** (mixed-integer convex optimization bound). *Let  $S = \mathbb{Z}^n \times \mathbb{R}^n$ , and  $\nu$  is the mixed integer measure on  $S$ . Then for any cutting plane algorithm  $\mathcal{A}$  there exists an instance such that  $\mathcal{A}$  makes at least  $2^n(\log_2(\frac{B^d}{\delta}) + n - 1)$  first-order evaluation oracle calls on  $\hat{g}$ .*

*Proof.* The proof is similar to the proof of Theorem 2.11. We can construct an adversarial function, that treats each fiber  $F_v := \{v\} \times [0, B]$ , where  $v \in \{0, 1\}^n$ , as a separate  $d$ -dimensional continuous problem.

For all fibers  $F_v$ , let  $\delta_v$  denote the measure of  $E_N$  intersected with  $F_v$ . By the stopping criteria, it must hold that  $\sum_v \delta_v \leq \delta$ . By Theorem 2.6 we know that at least  $(\log_2(\frac{B^d}{\delta_v})) - 1$  function oracle calls must be performed on each  $F_v$ , and by choosing our subgradient half-spaces in the same way as in the proof of Theorem 2.11, we obtain that the algorithm must make at least

$$N \geq \sum_{v \in \{0, 1\}^n} \left( \left( \log_2 \left( \frac{B^d}{\delta_v} \right) \right) - 1 \right) = 2^n(\log_2(B^d) - 1) - \log_2 \left( \prod_{v \in \{0, 1\}^n} \delta_v \right)$$

function oracle calls. Note that the last summand is minimized when  $\delta_v = \frac{\delta}{2^n}$  for all

$i = 1, \dots, 2^n$ . Hence,

$$N \geq 2^n(\log_2(B^d) - 1) - \log_2 \left( \left( \frac{\delta}{2^n} \right)^{2^n} \right) = 2^n \left( \log_2 \left( \frac{B^d}{\delta} \right) + n - 1 \right).$$

This completes the proof.  $\square$

We finish this section with a few remarks. As it was already proven by Nemirovsky and Yudin [35], the centerpoint algorithm is optimal for the continuous case up to a constant factor. For the pure integer case we could prove that our algorithm is optimal up to a linear factor in  $n$ . For the mixed-integer case, if Conjecture 3.5 would be true, we would have an upper bound of

$$\left\lceil \log_{\frac{e}{e-\frac{1}{2^n}}} \left( \frac{B^{n+d}}{\delta} \right) \right\rceil.$$

In particular, this would imply that the cutting plane algorithm, using centerpoints, is optimal for mixed-integer optimization, up to a linear factor only in  $n$ , which would nicely unify the continuous and discrete optimization theory; see Table 1.

Next, we want to point out that it is not difficult to generalize the cutting plane algorithm to the constrained optimization case:

$$\min_{\substack{x \in \mathbb{Z}^n \times \mathbb{R}^d, \\ h(x) \leq 0}} g(x),$$

where  $g, h : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}$  are convex functions given by first-order oracles. However, it is crucial that the feasible domain has a reasonably sized measure, as otherwise it might be impossible to find any feasible point, let alone an approximate optimal point. Further, the algorithm can be extended to handle quasi-convex functions, if one has access to separation oracles for their sublevel sets.

Finally, note that for the purely discrete case, when  $S = \mathbb{Z}^n$ , we can guarantee to find the optimal point of (4), provided we choose  $\delta < 1$ . Only when there are continuous variables, do we need to talk about approximations. Thus, let  $S = \mathbb{Z}^n \times \mathbb{R}^d$  with  $d \neq 0$ . We assume that for every fixed  $x \in \mathbb{Z}^n$ ,  $g(x, y)$  is Lipschitz continuous in the  $y$  variables with Lipschitz constant  $L$ . Let  $(\hat{x}, \hat{y}) \in \mathbb{Z}^n \times \mathbb{R}^d$  attain the optimal value  $\hat{g}$  of problem (4) and let  $(x^*, y^*)$  be the best point that the cutting plane algorithm has returned with objective value  $g^*$ . By standard arguments, we can bound  $\bar{\mu}_{mixed}(E_k)$  from below as follows:

$$\begin{aligned} \bar{\mu}_{mixed}(E_k) &\geq \bar{\mu}_{mixed}(\{(x, y) \in \mathbb{Z}^n \times \mathbb{R}^d : g((x, y)) - \hat{g} \leq g^* - \hat{g}\}) \\ &\geq \bar{\mu}_{mixed} \left( \left\{ (\hat{x}, y) \in \{\hat{x}\} \times \mathbb{R}^d : \|(\hat{x}, \hat{y}) - (\hat{x}, y)\|_2 \leq \frac{g^* - \hat{g}}{L} \right\} \right) \\ &= \left( \frac{\hat{g} - g^*}{L} \right)^d \kappa_d, \end{aligned}$$

where  $\kappa_d$  denotes the volume of the  $d$ -dimensional unit ball. On the other hand, after  $N$  iterations it holds that

$$\mu_{mixed}(E_N) \leq \delta.$$

Thus, we can guarantee that the algorithm returns a point satisfying  $g(x^*, y^*) - g(\hat{x}, \hat{y}) \leq L \left( \frac{\delta}{\kappa_d} \right)^{\frac{1}{d}}$ .

**3. Bounds on  $\mathcal{F}(S, \mu)$ .** We first establish some analytic properties of  $f_\mu$ . This will justify the use of “maximum” in (2), instead of a supremum. The goal of this section is to establish a bound on the quality of the centerpoints based on Helly numbers of  $S$ , which will be followed by a better lower bound when  $S$  is the mixed-integer lattice. We will denote the complement of a set  $X$  by  $X^c$ . We begin with a useful lemma.

**LEMMA 3.1.** *For any probability measure  $\mu$ ,  $f_\mu(x)$  defined in (3) is quasi-concave on  $\mathbb{R}^n$  and upper semicontinuous.*

*Proof.* For quasi-concavity, see [40, Proposition 1], and for upper semicontinuity, see [40, Proposition 4].  $\square$

**Remark 3.2.** Lemma 3.1 shows that  $\sup_{x \in S} f_\mu(x)$  is always attained. See [40, Proposition 7] where this is discussed for  $S = \mathbb{R}^n$ . The generalization to any nonempty, closed subset  $S$  is easy; see also [40, Proposition 5] which states the for every  $\alpha > 0$ , the set  $\{x \in \mathbb{R}^n : f_\mu(x) \geq \alpha\}$  is compact.

**3.1. A general lower bound based on Helly numbers.** We generalize a theorem well known in the literature on half-space (Tukey) depth [40, Proposition 9]; this was earlier stated by Grünbaum [22, Theorem 1] for uniform probability measures on convex bodies. In all of these works, the authors consider  $S = \mathbb{R}^n$ , as discussed in the introduction. We consider more general sets  $S$ . For this we define the Helly number of a set  $S \subseteq \mathbb{R}^n$ . Let  $\mathcal{K} := \{S \cap K \mid K \subseteq \mathbb{R}^n \text{ convex}\}$ . Then the Helly number  $h = h(S) \in \mathbb{N}$  of  $S$  is defined as the smallest number such that the following property is satisfied for all finite subsets  $\{C_1, \dots, C_m\} \subseteq \mathcal{K}$ : If

$$C_{i_1} \cap \dots \cap C_{i_h} \neq \emptyset \text{ for all } \{i_1, \dots, i_h\} \subseteq \{1, \dots, m\}$$

then

$$C_1 \cap \dots \cap C_m \neq \emptyset.$$

If no such number exists, then  $h(S) = \infty$ . This extension of Helly’s number was first considered by Hoffman [26], and has recently been studied in [2, 1, 15].

**THEOREM 3.3.** *Let  $S \subseteq \mathbb{R}^n$  be a nonempty, closed subset and let  $\mu$  be such that  $\mu(\mathbb{R}^n \setminus S) = 0$ . If  $h(S) < \infty$ , then  $\mathcal{F}(S, \mu) \geq h(S)^{-1}$ .*

*Proof.* The proof follows along similar lines to [40, Proposition 9]. It suffices to show that for every  $\epsilon > 0$ , the set  $\{x \in \mathbb{R}^n : f_\mu(x) \leq h(S)^{-1} - \epsilon\}$  is nonempty. By standard measure-theoretic arguments, there exists a ball  $B$  centered at the origin such that  $\mu(B) \geq 1 - \frac{\epsilon}{2}$  and  $\{x \in \mathbb{R}^n : f_\mu(x) \leq h(S)^{-1} - \epsilon\} \subseteq B$  (by Remark 3.2,  $\{x \in \mathbb{R}^n : f_\mu(x) \leq h(S)^{-1} - \epsilon\}$  is compact). By [40, Proposition 6],

$$\begin{aligned} & \{x \in S : f_\mu(x) \leq h(S)^{-1} - \epsilon\} \\ &= \bigcap \{H \cap S : H \text{ is a closed half-space with } \mu(H) \geq 1 - (h(S)^{-1} - \epsilon)\}. \end{aligned}$$

Define  $\mathcal{C} = \{B \cap H \cap S : H \text{ is a closed half-space with } \mu(H) \geq 1 - (h(S)^{-1} - \epsilon)\}$ . Thus,  $\mathcal{C}$  is a family of compact sets such that  $\{x \in S : f_\mu(x) \leq h(S)^{-1} - \epsilon\} = \bigcap \{C : C \in \mathcal{C}\}$ . For any subset  $\{C_1, \dots, C_{h(S)}\} \subseteq \mathcal{C}$  of size  $h(S)$ , we claim

$$\mu(C_1^c \cup \dots \cup C_{h(S)}^c) \leq 1 - h(S)^{-1} + \frac{\epsilon}{2}.$$

This is because each  $C_i^c = B^c \cup H_i^c \cup S^c$  for some half-space  $H_i$  satisfying  $\mu(H_i^c) \leq h(S)^{-1} - \epsilon$ . Since  $\mu(B^c) \leq \frac{\epsilon}{2}$  and  $\mu(S^c) = 0$ , we obtain that  $\mu(C_i^c) \leq h(S)^{-1} - \frac{\epsilon}{2}$ .

Therefore,

$$\mu(C_1 \cap \dots \cap C_{h(S)}) = 1 - (\mu(C_1^c \cup \dots \cup C_{h(S)}^c)) \geq 1 - \left(1 - h(S) \frac{\epsilon}{2}\right) = h(S) \frac{\epsilon}{2} > 0.$$

This implies that  $C_1 \cap \dots \cap C_{h(S)} \neq \emptyset$ . Therefore, by the definition of  $h(S)$ , for every finite subset  $\{C_1, \dots, C_m\} \subseteq \mathcal{C}$ ,  $C_1 \cap \dots \cap C_m \neq \emptyset$ . By the finite intersection property of compact sets, we obtain that  $\{x \in S : f_\mu(x) \leq h(S)^{-1} - \epsilon\} = \bigcap \{C : C \in \mathcal{C}\}$  is nonempty.  $\square$

By applying the well-known bound for the mixed-integer Helly number [26, 2, 15] we get the following corollary.

**COROLLARY 3.4.**  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \mu) \geq \frac{1}{2^n(d+1)}$  for any finite measure  $\mu$  on  $\mathbb{R}^{n+d}$  such that  $\mu(\mathbb{R}^{n+d} \setminus (\mathbb{Z}^n \times \mathbb{R}^d)) = 0$ . In particular, this holds for  $\mu_{\text{mixed}, K}$  for any convex body  $K \subseteq \mathbb{R}^n \times \mathbb{R}^d$ .

**3.2. Better bounds for the mixed-integer lattice.** We would like to improve the bound on  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu)$  coming from Helly numbers (Theorem 3.3 and Corollary 3.4) when  $\nu$  is a mixed-integer measure. Better bounds have been obtained by Grünbaum for the purely continuous case ( $n = 0$ ), by exploiting properties of the *centroid* of a convex body  $K$ , which is defined as  $c_K := \int_K x \, d\mu(x)$ , where the integral is taken with respect to the uniform measure  $\mu$  on  $K$ . Grünbaum proved in [22] that  $\mu(H^+(u, c_K)) \geq (\frac{d}{d+1})^d \geq e^{-1}$  for any  $u \in S^{d-1}$ , which immediately implies that  $\mathcal{F}(\mathbb{R}^d, \mu) \geq e^{-1}$ . This, of course, drastically improves the Helly bound of  $\frac{1}{d+1}$ . Note that, even though the centerpoint and centroid are equal for several extreme cases, this is, in general, not true. In the following we want to extend these improved bounds to the mixed-integer setting. Ideally, we would want to prove the following conjecture.

**CONJECTURE 3.5.** Let  $S = \mathbb{Z}^n \times \mathbb{R}^d$  and let  $\nu := \mu_{\text{mixed}, K}$  for some convex body  $K \in \mathbb{R}^{n+d}$ . Then  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu) \geq \frac{1}{2^n} (\frac{d}{d+1})^d \geq \frac{1}{2^n} \frac{1}{e}$ .

While we have not been able to resolve the above conjecture, we show that it holds in the regime of convex sets  $K$  with “large” *lattice width*, where the lattice width is defined as

$$\omega(K) := \min_{z \in \mathbb{Z}^n \setminus \{0\}} \left[ \max_{x \in K} u \cdot x - \min_{x \in K} u \cdot x \right].$$

In fact, we prove something stronger in this regime.

**THEOREM 3.6.** There exists a universal constant  $\alpha$  such that for all  $n, d \in \mathbb{N}$  and any convex body  $K \subseteq \mathbb{R}^{n+d}$  with  $\omega(K) > 2cn(n+d)^{5/2}\alpha n^{n+1}$  for some  $c \in \mathbb{R}_+$ , the following holds:

$$\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu) \geq e^{-\frac{1}{c}-1} + e^{-\frac{2}{c}} - 1.$$

In particular, when  $c \in \mathbb{R}_+$  is such that  $e^{-\frac{1}{c}-1} + e^{-\frac{2}{c}} - 1 \geq 2^{-n-1}$ , we have  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu) \geq \frac{1}{2^n} \frac{1}{2} \geq \frac{1}{2^n} (\frac{d}{d+1})^d \geq \frac{1}{2^n} \frac{1}{e}$ .

We denote the projection of a set  $X \subseteq \mathbb{R}^{n+d}$  onto the first  $n$  coordinates by  $X|_{\mathbb{R}^n}$ .

**Remark 3.7.** Theorem 3.6 provides some evidence towards our belief in Conjecture 3.5. In particular, we see that it holds in two distinct regimes.

Suppose the convex set  $K$  is such that  $K|_{\mathbb{R}^n}$  is “thin” in every direction; more precisely, suppose there exists a constant  $C$  such that for every unit vector  $e^i$ ,  $i =$

$1, \dots, n$ , we have  $\max_{x \in K} e^i \cdot x - \min_{x \in K} e^i \cdot x \leq C$ . Then  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu) \geq \frac{1}{C^n} \left(\frac{d}{d+1}\right)^d$ . This is witnessed by choosing the centroid of the fiber with at least  $\frac{1}{C^n}$  fraction of the total mass of  $K \cap (\mathbb{Z}^n \times \mathbb{R}^d)$ —such a fiber exists because there are at most  $C^n$  fibers intersecting  $K$ .

On the other hand, suppose  $K$  is such that  $K|_{\mathbb{R}^n}$  is “fat” in every direction; more precisely, the hypothesis of Theorem 3.6 holds. Then we get an even stronger bound than  $\frac{1}{2^n} \left(\frac{d}{d+1}\right)^d$  from Theorem 3.6.

The rest of this section is devoted to the proof of Theorem 3.6. The main ingredient in the proof of Theorem 3.6 is Lemma 3.8, where we show that for convex sets with large lattice width, the  $d$ -dimensional Lebesgue measure  $\bar{\nu} := \bar{\mu}_{mixed}$  of  $K \cap (\mathbb{Z}^n \times \mathbb{R}^d)$  can be approximated by the  $(d+n)$ -dimensional Lebesgue measure  $\bar{\mu}$  of  $K$  and vice versa. (Note that in this case we do not normalize the measures.) In the pure integer setting, i.e.,  $d = 0$ , this connection is well known. However, to the best of our knowledge, this kind of result has never been proven for the mixed-integer setting nor explicitly with respect to the lattice width.

**LEMMA 3.8.** *There exists a universal constant  $\alpha$  such that for all  $n, d \in \mathbb{N}$  and any convex body  $K \subseteq \mathbb{R}^{n+d}$  and  $\omega(K|_{\mathbb{R}^n}) \geq cn(n+d)^{5/2}\alpha n^{n+1}$  for some  $c \in \mathbb{R}_+$ , then the following holds:*

$$e^{-\frac{1}{c}} \leq \frac{\bar{\nu}(K \cap (\mathbb{Z}^n \times \mathbb{R}^d))}{\bar{\mu}(K)} \leq e^{\frac{1}{c}}.$$

For the proof of Lemma 3.8 we need two technical auxiliary lemmas. The first lemma, Lemma 3.9, gives an ellipsoidal approximation of a convex body using the centerpoint as the center of the two ellipsoids used for the approximation. This is a variation on the classical Fritz-John ellipsoidal approximation result. In the second lemma, Lemma 3.10, we show that for a convex body  $K$  with large lattice-width, there exists a basis for the mixed-integer lattice such that  $K$  contains a scaled copy of the fundamental parallelepiped of the lattice with respect to this basis, centered at the centerpoint of  $K$ .

**LEMMA 3.9.** *Let  $K \subseteq \mathbb{R}^n$  be a compact convex set with nonempty interior and let  $\mu$  be the uniform measure with respect to  $K$ . Further let  $x^* \in \mathcal{C}(\mathbb{R}^n, \mu)$ . Then, there exists an ellipsoid  $E$  centered at the origin such that*

$$x^* + E \subseteq K \subseteq x^* + n^{5/2}E.$$

*Proof.* Without loss of generality we assume that  $x^* = 0$ . We prove that for any  $u \in S^{n-1}$

$$(8) \quad \frac{1}{n^2} \leq \left| \frac{\max_{x \in K} u^\top x}{\min_{x \in K} u^\top x} \right| \leq n^2.$$

It suffices to show the lower bound, since the upper bound follows from replacing  $u$  with  $-u$ .

Let  $\left| \frac{\max_{x \in K} u^\top x}{\min_{x \in K} u^\top x} \right|$  be minimized at  $u = \bar{u}$ . Since the arguments below are invariant to scaling, we will assume that  $\min_{x \in K} \bar{u}^\top x = -1$  and  $\max_{x \in K} \bar{u}^\top x =: \alpha$ , and assume to the contrary that  $\alpha < \frac{1}{n^2}$ . Let  $z := \operatorname{argmin}_{x \in K} \bar{u}^\top x$ . We define for every  $t \in \mathbb{R}$  the set  $K_t := K \cap \{x \in \mathbb{R}^n : \bar{u}^\top x = t\}$ . Further, we define  $C := z + \operatorname{cone}(K_0 - z)$ ,  $X_1 := \{x \in \mathbb{R}^n : -1 \leq \bar{u}^\top x \leq 0\}$ , and  $X_2 = \{x \in \mathbb{R}^n : 0 \leq \bar{u}^\top x \leq \alpha\}$ . Then  $K \cap X_1 \supset C \cap X_1$  and  $K \cap X_2 \subseteq C \cap X_2$ . By Grünbaum’s theorem [22, Theorem 2] we have that  $1 - \left(\frac{n}{n+1}\right)^n \geq \mu(K \cap X_1) \geq \mu(C \cap X_1) = \frac{1}{n} \frac{V}{\operatorname{vol}(K)}$ , where  $V$  represents the

$(n-1)$ -dimensional measure of  $K_0$ . On the other hand we have  $(\frac{n}{n+1})^n \leq \mu(K \cap X_2) \leq \mu(C \cap X_2) = (\frac{1+\alpha}{n} - \frac{1}{n}) \frac{V}{\text{vol}(K)}$ . Combining these two inequalities, we arrive at the inequality

$$\frac{1}{e-1} \leq \frac{(\frac{n}{n+1})^n}{1 - (\frac{n}{n+1})^n} \leq (1+\alpha)^n - 1 \leq \left(1 + \frac{1}{n^2}\right)^n - 1.$$

However,  $(1 + \frac{1}{n^2})^n - 1 < \frac{1}{e-1}$  for all  $n \geq 2$ , leading to a contradiction.

We now define

$$\text{sym}(K) = \max\{\alpha \geq 0 : \alpha(-y) \in K \text{ for every } y \in K\},$$

a notion that was introduced by Minkowski, and has been extensively studied in convex geometry literature [6, 42]. Combined with [6, Proposition 1], (8) implies that  $\text{sym}(K) \geq \frac{1}{n^2}$ . Then [6, Theorem 7] shows that there exists an ellipsoid  $E$  centered at the origin satisfying  $E \subseteq K \subseteq n^{5/2}E$ .  $\square$

LEMMA 3.10. *There exists a universal constant  $\alpha$  such that the following holds for all  $n, d \in \mathbb{N}$ . Let  $K \subseteq \mathbb{R}^{n+d}$  be a convex body and let  $x^* \in \mathcal{C}(\mathbb{R}^n, \mu)$ , where  $\mu$  is the uniform measure with respect to  $K$ . If  $\omega(K|_{\mathbb{R}^n}) \geq cn(n+d)^{5/2}\alpha n^{n+1}$  for some  $c \in \mathbb{R}_+$ , then there exists a matrix  $B = [b_1, \dots, b_n] \in \mathbb{R}^{(n+d) \times n}$  such that*

$$x^* + cnB[-1/2, 1/2]^n \subseteq K,$$

and  $b_1|_{\mathbb{R}^n}, \dots, b_n|_{\mathbb{R}^n}$  forms a lattice basis of  $\mathbb{Z}^n$ .

*Proof.* By Lemma 3.9, there exists an ellipsoid  $E$  such that  $x^* + E \subseteq K \subseteq x^* + (n+d)^{5/2}E$ . We define  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  as the linear map such that  $\phi(E|_{\mathbb{R}^n}) = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ . Let  $\Lambda := \phi(\mathbb{Z}^n)$ .

Let  $\hat{B}$  be a matrix whose columns  $\hat{B}_{*,1}, \dots, \hat{B}_{*,n}$  form a Korkine–Zolotarev basis of  $\Lambda$  [28]. Then, a well-known property is that

$$(9) \quad \|\hat{B}_{*,1}\|_2 \cdots \|\hat{B}_{*,n}\|_2 \leq \alpha n^n \det(\Lambda)$$

(see [30, Theorem 2.3]), where  $\alpha$  is a universal constant. Further, the Gram–Schmidt orthogonalization  $\tilde{B}_{*,1}, \dots, \tilde{B}_{*,n}$  of  $\hat{B}$  satisfies

$$(10) \quad \|\tilde{B}_{*,i}\|_2 \leq \|\hat{B}_{*,i}\|_2 \text{ for all } i = 1, \dots, n$$

and it holds that

$$(11) \quad \det(\Lambda) = \det(\hat{B}) = \det(\tilde{B}) = \prod_{i=1}^n \|\tilde{B}_{*,i}\|_2$$

(see, for example, [21, Chapter 28]). Since  $K \subseteq x^* + (n+d)^{5/2}E$ , the definition of the lattice width  $\omega(K|_{\mathbb{R}^n})$  implies that

$$(12) \quad \|\tilde{B}_{*,n}\|_2 \leq \frac{2(n+d)^{5/2}}{\omega(K|_{\mathbb{R}^n})}.$$

Putting all these relations together, for all  $j \in \{1, \dots, n\}$ , we have the following:

$$\begin{aligned} (\prod_{i \neq n} \|\tilde{B}_{*,i}\|) \cdot \|\hat{B}_{*,n}\| &\leq (\prod_{i \neq n} \|\hat{B}_{*,i}\|) \cdot \|\hat{B}_{*,n}\| && \text{using (10)} \\ &\leq \alpha n^n \det(\Lambda) && \text{using (9)} \\ &= \alpha n^n (\prod_{i=1}^n \|\tilde{B}_{*,i}\|) && \text{using (11)} \\ &\leq (\prod_{i \neq n} \|\tilde{B}_{*,i}\|) \alpha n^n \frac{2(n+d)^{5/2}}{\omega(K|_{\mathbb{R}^n})} && \text{using (12)} \end{aligned}$$

and, thus, we obtain

$$\|\widehat{B}_{\star,n}\|_2 \leq \frac{2\alpha n^n (n+d)^{5/2}}{\omega(K|_{\mathbb{R}^n})}.$$

If we change the order of the columns in  $\widehat{B}$ , (9)–(12) still hold (with a different  $\widehat{B}_{\star,n}$ ), and thus we obtain a bound on the Euclidean length of all Korkine–Zolotarev vectors, i.e., for all  $i = 1, \dots, n$

$$\|\widehat{B}_{\star,i}\|_2 \leq \frac{2\alpha n^n (n+d)^{5/2}}{\omega(K|_{\mathbb{R}^n})}.$$

This implies that

$$\frac{1}{n} \frac{1}{\alpha n^n} \frac{\omega(K|_{\mathbb{R}^n})}{(n+d)^{5/2}} \widehat{B}[-1/2, 1/2]^n \subseteq \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}.$$

Since  $\omega(K|_{\mathbb{R}^n}) \geq cn(n+d)^{5/2}\alpha n^{n+1}$ , we obtain that

$$cn\widehat{B}[-1/2, 1/2]^n \subseteq \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}.$$

We now use the fact that there exists an affine subspace  $H \subseteq \mathbb{R}^{n+d}$  of dimension  $n$  such that  $E|_{\mathbb{R}^n} = (E \cap H)|_{\mathbb{R}^n}$ . This follows from a repeated application of [16, Lemma 3]. Let  $B$  be the matrix whose columns span the linear space parallel to  $H$  and these columns project to  $\phi^{-1}(\widehat{B}_{\star,1}), \dots, \phi^{-1}(\widehat{B}_{\star,n})$ .  $B$  now satisfies the condition desired.  $\square$

We are now ready to prove Lemma 3.8.

*Proof of Lemma 3.8.* By Lemma 3.10 there exists a matrix  $B \in \mathbb{R}^{(n+d) \times n}$  such that  $x^* + cnB[-1/2, 1/2]^n \subseteq K$ . Since  $B|_{\mathbb{R}^n}$  is unimodular, we may assume after a unimodular transformation that  $B|_{\mathbb{R}^n}$  equals the identity matrix. After a further volume preserving linear transformation we may even assume that  $B$  equals the first  $n$  unit vectors in  $\mathbb{R}^{n+d}$ . Since  $K$  is full dimensional, there exists an  $\epsilon > 0$  such that  $cn[-1/2, 1/2]^n \times \epsilon[-1/2, 1/2]^d \subseteq K - x^*$ . To make the calculations below easy, we translate everything by  $-x^*$ , so that we assume below that  $x^* = 0$ .

We next exploit that  $\lim_{k \rightarrow \infty} \frac{1}{k^d} |K \cap (\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d)| = \bar{\nu}(K \cap (\mathbb{Z}^n \times \mathbb{R}^d))$ , where  $|\cdot|$  denotes the cardinality. We will use this fact to prove the following.

CLAIM 3.11.

$$\left(1 - \frac{1}{cn}\right)^n \left(1 - \frac{1}{\epsilon k}\right)^d \leq \frac{\frac{1}{k^d} |K \cap (\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d)|}{\bar{\mu}(K)} \leq \left(1 + \frac{1}{cn}\right)^n \left(1 + \frac{1}{\epsilon k}\right)^d.$$

*Proof.* Let  $Q := [-1/2, 1/2]^n \times [-1/2k, 1/2k]^d$ . Further we define

$$\bar{K} := \left(K \cap \left(\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d\right)\right) + Q$$

and  $K_{\lambda,\gamma} := \begin{pmatrix} \lambda I_n & 0 \\ 0 & \gamma I_d \end{pmatrix} K$  for  $\lambda, \gamma \in \mathbb{R}_+$ . Notice that  $K_{\lambda_1,\gamma_1} + K_{\lambda_2,\gamma_2} = K_{\lambda_1+\lambda_2,\gamma_1+\gamma_2}$  and  $Q \subseteq K_{\frac{1}{cn}, \frac{1}{\epsilon k}}$ .

Since  $\frac{1}{k^d} |K \cap (\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d)| = \bar{\mu}(\bar{K})$ , it suffices to show that

$$K_{1-\frac{1}{cn}, 1-\frac{1}{\epsilon k}} \subseteq \bar{K} \subseteq K_{1+\frac{1}{cn}, 1+\frac{1}{\epsilon k}}.$$

One of the containments follows from the observations that  $\bar{K} \subseteq K + Q$  and  $K + Q \subseteq K_{1+\frac{1}{cn}, 1+\frac{1}{\epsilon k}}$ . In order to prove  $K_{1-\frac{1}{cn}, 1-\frac{1}{\epsilon k}} \subseteq \bar{K}$ , suppose to the contrary that there



exists an  $x \in K_{1-\frac{1}{cn}, 1-\frac{1}{ek}} \setminus \bar{K}$ . We define  $z \in \mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d$ , such that  $z_i = \lfloor x_i \rfloor$  for  $1 \leq i \leq n$  and  $z_i = \frac{1}{k} \lfloor kx_i \rfloor$  for  $n < i \leq n+d$ .<sup>2</sup> Then, since  $z - x \in Q \subseteq K_{\frac{1}{cn}, \frac{1}{ek}}$  and  $x \in K_{1-\frac{1}{cn}, 1-\frac{1}{ek}}$ , we obtain that  $z$  must be in  $K \cap (\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d)$ . Since  $Q$  is symmetric,  $x - z$  is also in  $Q$ , and therefore  $x \in z + Q \subseteq K \cap (\mathbb{Z}^n \times \frac{1}{k}\mathbb{Z}^d) + Q = \bar{K}$ , which contradicts the choice of  $x$ .  $\square$

We get the desired inequalities from Claim 3.11 by letting  $k$  go to infinity and using the fact that  $(1 + \frac{1}{cn})^n \leq e^{\frac{1}{c}}$  and  $(1 - \frac{1}{cn})^n \geq e^{-\frac{1}{c}}$  for all  $c \in \mathbb{R}_+$  and  $n \in \mathbb{N}$ .  $\square$

To complete the proof of Theorem 3.6 we introduce the following technical rounding procedure. Let  $K$  be a convex body with a sufficiently large lattice width, i.e.,  $\omega(K) > cn(n+d)^{5/2}\alpha n^n \sqrt{n}$  for some positive integer  $c$ , where  $\alpha$  is the constant from Lemma 3.8. Let  $\mu$  be the uniform measure on  $K$  and let  $x^* \in \mathcal{C}(\mathbb{R}^{n+d}, \mu)$ . By Lemma 3.10, there exist  $b_i \in (-x^* + K) \cap (\mathbb{Z}^n \times \mathbb{R}^d)$  for  $i = 1, \dots, n$  such that  $b_1|_{\mathbb{R}^n}, \dots, b_n|_{\mathbb{R}^n}$  is a lattice basis of  $\mathbb{Z}^n$  and  $x^* + cnB[-1/2, 1/2]^n \subseteq K$ . In addition we define  $b_i \in \mathbb{R}^{n+d}$  as the  $i$ th unit vector for  $i = n+1, \dots, n+d$ . Hence,  $b_1, \dots, b_{n+d}$  define a basis of  $\mathbb{R}^{n+d}$ .

Given  $x = \sum_{i=1}^{n+d} \lambda_i b_i \in \mathbb{R}^{n+d}$  with  $\lambda_i \in \mathbb{R}$  for all  $i$ , we define  $[x]_K \in \mathbb{Z}^n \times \mathbb{R}^d$  as  $\sum_{i=1}^n \lfloor \lambda_i \rfloor b_i + \sum_{i=n+1}^{n+d} \lambda_i b_i$ , i.e., we round  $x$  to a close mixed-integer point with respect to  $K$  (the dependence on  $K$  is through Lemma 3.10 which defines the matrix  $B$ ).

**THEOREM 3.12.** *Let  $\alpha$  be the constant from Lemma 3.8. Let  $\nu := \mu_{\text{mixed}, K}$ , where  $K \subseteq \mathbb{R}^{n+d}$  is a convex body and  $\nu(K) \neq 0$ , and let  $x^*$  be the centerpoint with respect to  $\mu$ , the uniform measure on  $K$ . If  $\omega(K|_{\mathbb{R}^n}) > 2cn(n+d)^{5/2}\alpha n^{n+1}$  for some  $c \in \mathbb{R}_+$ , then*

$$f_\nu([x^*]_K) \geq e^{-\frac{1}{c}} \mathcal{F}(\mathbb{R}^{d+n}, \mu) + e^{-\frac{2}{c}} - 1.$$

Grünbaum's theorem implies then, that  $\mathcal{F}(\mathbb{Z}^n \times \mathbb{R}^d, \nu) \geq e^{-\frac{1}{c}-1} + e^{-\frac{2}{c}} - 1$ , giving us Theorem 3.6.

*Proof of Theorem 3.12.* As before, let  $\bar{\mu}$  denote the  $(d+n)$ -dimensional Lebesgue measure with respect to  $K$  and let  $\bar{\nu}$  denote the  $d$ -dimensional Lebesgue measure with respect to  $K \cap (\mathbb{Z}^n \times \mathbb{R}^d)$ , i.e., they are not normalized.

In a first step we prove the following claim.

**CLAIM 3.13.** *For any half-space  $H^+$ ,*

$$\nu(H^+) \geq \mu(H^+) + e^{-\frac{2}{c}} - 1.$$

*Proof of Claim 3.13.* Let  $H^+$  be any half-space and let  $H^-$  denote its closed complement. The lattice width of either  $(K \cap H^+)|_{\mathbb{R}^n}$  or  $(K \cap H^-)|_{\mathbb{R}^n}$  is larger than or equal to  $\omega(K|_{\mathbb{R}^n})/2$ . If  $\omega(K \cap H^-) \geq cn(n+d)^{5/2}\alpha n^n \sqrt{n}$ , then, by Lemma 3.8,

$$\begin{aligned} \nu(H^+) &= \frac{\bar{\nu}(K \cap H^+)}{\bar{\nu}(K)} \geq \frac{e^{-\frac{1}{c}} \bar{\mu}(K) - e^{\frac{1}{c}} \bar{\mu}(K \cap H^-)}{e^{\frac{1}{c}} \bar{\mu}(K)} \\ &= \frac{\bar{\mu}(K \cap H^+)}{\bar{\mu}(K)} + \frac{(e^{-\frac{1}{c}} - e^{\frac{1}{c}}) \bar{\mu}(K)}{e^{\frac{1}{c}} \bar{\mu}(K)} \\ &= \mu(H^+) + e^{-\frac{2}{c}} - 1. \end{aligned}$$

<sup>2</sup>For  $x \in \mathbb{R}$ ,  $\lfloor x \rfloor$  denotes the integer  $z \in \mathbb{Z}$  such that for each component  $|z - x| \leq \frac{1}{2}$ .

If  $\omega(K \cap H^+) \geq cn(n+d)\alpha n^n \sqrt{n}$ , then, by Lemma 3.8,

$$\begin{aligned}\nu(H^+) &= \frac{\bar{\nu}(K \cap H^+)}{\bar{\nu}(K)} \geq \frac{e^{-\frac{1}{c}} \bar{\mu}(K \cap H^+)}{e^{\frac{1}{c}} \bar{\mu}(K)} \\ &= e^{-\frac{2}{c}} \mu(H^+) \\ &= \mu(H^+) - (1 - e^{-\frac{2}{c}}) \mu(H^+) \\ &\geq \mu(H^+) + e^{-\frac{2}{c}} - 1.\end{aligned}$$

The last inequality holds since  $\mu(H^+) \leq 1$  and  $1 - e^{-\frac{2}{c}} \geq 0$ .  $\square$

In the second step we bound the error made by rounding the  $x^*$  to  $[x^*]_K$ . Note that this is done with respect to a matrix  $B$  that is returned from Lemma 3.10. We first make a unimodular transformation so that  $B|_{\mathbb{R}^n}$  is the identity and then make an affine transformation so that  $B$  consists of the first  $n$  unit vectors in  $\mathbb{R}^{n+d}$ . To make a further simplification, we translate everything by  $-x^*$  so that  $x^* = 0$ .

Define  $K_{\lambda, \gamma} := \begin{pmatrix} \lambda I_n & 0 \\ 0 & \gamma I_d \end{pmatrix} K$  for any  $\lambda, \gamma \geq 0$ . Since we assume  $x^* = 0$ , our rounding procedure implies that  $[x^*] \in B[-1/2, 1/2]^n$ . By Lemma 3.10, this implies that  $[x^*] \in K_{\frac{1}{cn}, 0}$ . Therefore,  $[x^*] + K_{1-\frac{1}{cn}, 1} \subseteq K$ . This implies that for any  $u \in \mathcal{S}^{n+d-1}$ ,  $\mu(H^+(u, [x^*]_K)) \geq (1 - \frac{1}{cn})^n \mu(H^+(u, x^*)) \geq e^{-\frac{1}{c}} \mu(H^+(u, x^*))$ .

Together with the previous claim it follows that

$$\begin{aligned}f_\nu([x^*]_K) &= \max_{u \in \mathcal{S}^{n+d-1}} \nu(H^+(u, [x^*]_K)) \\ &\geq \max_{u \in \mathcal{S}^{n+d-1}} \mu(H^+(u, [x^*]_K)) + e^{-\frac{2}{c}} - 1 \\ &\geq \max_{u \in \mathcal{S}^{n+d-1}} e^{-\frac{1}{c}} \mu(H^+(u, x^*)) + e^{-\frac{2}{c}} - 1 \\ &\geq e^{-\frac{1}{c}} \mathcal{F}(\mathbb{R}^{d+n}, \mu) + e^{-\frac{2}{c}} - 1.\end{aligned}$$

This completes the proof.  $\square$

**4. Computational aspects.** All our algorithms discussed in this section are under the standard Turing machine model of computation. We say that  $x \in S$  is an  $\epsilon$ -centerpoint for  $S, \mu$ , if  $f_\mu(x) \geq \mathcal{F}(S, \mu) - \epsilon$ , where  $\mathcal{F}(S, \mu)$  is defined in (2) and  $f_\mu$  is defined in (3).

A central tool in the following algorithms for computing (approximate) centerpoints is solving convex mixed-integer optimization problems. The classical result here is due to Grötschel, Lovász, and Schrijver [20]. This classical algorithm requires an access to a first-order oracle for the convex function at all points in  $\mathbb{R}^n$ . It can be modified to solve the problem with access to a first-order oracle that only queries mixed-integer points (as opposed to any point in  $\mathbb{R}^n$ ). This modification will be useful for us in this section, in particular, for Theorem 4.4. For completeness, we give a description of the result most amenable for our purposes; an appropriate reference for this version is [38].

**THEOREM 4.1.** *Let  $S = \mathbb{Z}^n \times \mathbb{R}^d$ ,  $B \geq 0$ , and  $\epsilon > 0$ . Let  $f : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}$  be a quasi-concave function equipped with an oracle such that there exists  $\delta \leq (C \frac{\epsilon}{B})^{n+d}$  for some universal constant  $C$  independent of  $B, n, d, \epsilon$  with the following property. For any point  $(\bar{x}, \bar{y}) \in S$ , the oracle returns an approximate function value  $\bar{f}$  and an approximate separation vector  $\bar{u} \in \mathcal{S}^{n+d-1}$  with the following guarantees:*

- (i) *There exists an optimal solution in  $\operatorname{argmax}_{x \in S} f(x)$  with norm at most  $B$ ,*
- (ii)  *$|f(\bar{x}, \bar{y}) - \bar{f}| \leq \delta$ ,*

- (iii)  $\|u - \bar{u}\|_\infty \leq \delta$  for some  $u \in \mathcal{S}^{n+d-1}$  satisfying  $\{(x, y) \in S : f(x, y) \geq f(\bar{x}, \bar{y})\} \subseteq \{(x, y) : u \cdot (x, y) \leq u \cdot (\bar{x}, \bar{y})\}$ .

Then there is an algorithm that computes a point  $x^* \in S$  such that  $\max_{x \in S} f(x) - f(x^*) \leq \epsilon$ . Moreover, if  $n$  is fixed, the algorithm runs in time polynomial in  $\log(B)$ ,  $\log(\frac{1}{\epsilon})$ , and  $d$ .

#### 4.1. Exact algorithms.

**4.1.1. Uniform measure on polytopes.** Since the rationality of the centerpoint for uniform measures on rational polytopes is an open question, we consider an “exact” algorithm as one which returns an  $\epsilon$ -centerpoint and runs in time polynomial in  $\log(\frac{1}{\epsilon})$  and the size of the description of the rational polytope.

**THEOREM 4.2.** *Let  $n$  be a fixed natural number. There is an algorithm which takes as input a rational polytope  $P \subseteq \mathbb{R}^n$  and  $\epsilon > 0$ , and returns an  $\epsilon$ -centerpoint for  $S = \mathbb{R}^n$  and  $\mu$ , the uniform measure on  $P$ . The algorithm runs in time polynomial in the size of an irredundant description of  $P$  and  $\log(\frac{1}{\epsilon})$ .*

*Proof.* Since  $f_\mu$  defined in (3) is quasi-concave by Lemma 3.1, an  $x^*$  satisfying  $f_\mu(x^*) \geq \mathcal{F}(S, \mu) - \epsilon$  can be found using Theorem 4.1, if one has an approximate evaluation oracle for  $f_\mu$ , and an approximate separation oracle for the level sets.

Implementing these oracles boils down to the following: Given  $\bar{x} \in \mathbb{R}^n$  and  $\delta > 0$ , find  $\bar{u} \in \mathcal{S}^{n-1}$  such that

$$(13) \quad \mu(H^+(\bar{u}, \bar{x})) \leq \min_{u \in \mathcal{S}^{n-1}} \mu(H^+(u, \bar{x})) + \delta \quad \text{and} \quad \|u - \bar{u}\|_\infty \leq \delta$$

for some  $u \in \operatorname{argmin}_{u \in \mathcal{S}^{n-1}} \mu(H^+(u, \bar{x}))$ .

Given  $\bar{x}$ , let  $\mathcal{P}$  be the set of all partitions of the vertices of  $P$  into two sets that can be achieved by a hyperplane through  $\bar{x}$ . (Note that, since  $n$  is assumed to be fixed, the number of vertices of  $P$  is polynomially bounded and they can be computed in time bounded by a polynomial in terms of the input size of  $P$ .) This induces a covering of the sphere  $\mathcal{S}^{n-1}$ : For each  $X \in \mathcal{P}$  define  $U_X$  to be the set of  $u \in \mathcal{S}^{n-1}$  such that the hyperplane  $u \cdot x = u \cdot \bar{x}$  induces the partition  $X$  on the vertices of  $P$ . The number of such partitions is closely related to the Vapnik–Chervonenkis (VC) dimension of hyperplanes and, in particular, is easily seen to be  $O(M^n)$ , where  $M$  is the number of vertices of  $P$ . Indeed, let  $X \in \mathcal{P}$ . Then there exists a  $u \in \operatorname{cl}(U_X)$  (where  $\operatorname{cl}(\cdot)$  denotes the closure), such that the hyperplane defined by  $u$  contains  $n - 1$  vertices  $\{v_1, \dots, v_{n-1}\}$  of  $P$  such that  $\{\bar{x}, v_1, \dots, v_{n-1}\}$  are affinely independent. Thus, we can construct a  $u_X \in U_X$  by perturbing this hyperplane to obtain the partition  $X$ . Moreover, one can enumerate these partitions in the same amount of time, by picking  $n - 1$  vertices  $\{v_1, \dots, v_{n-1}\}$  of  $P$  such that  $\{\bar{x}, v_1, \dots, v_{n-1}\}$  are affinely independent.

To solve problem (13), we will proceed along these steps.

1. For each  $X \in \mathcal{P}$ , find  $\bar{u}_X \in \mathcal{S}^{n-1}$  be such that

$$\mu(H^+(\bar{u}_X, \bar{x})) \leq \min_{u \in U_X} \mu(H^+(u, \bar{x})) + \delta \quad \text{and} \quad \|u - \bar{u}_X\|_\infty \leq \delta$$

for some  $u \in \operatorname{argmin}_{u \in U_X} \mu(H^+(u, \bar{x}))$ .

2. Pick  $X^*$  such that  $\mu(H^+(\bar{u}_{X^*}, \bar{x})) \leq \mu(H^+(\bar{u}_X, \bar{x}))$  for all  $X \in \mathcal{P}$  and report  $\bar{u}_{X^*}$  as the solution to (13).

To complete the proof, we need to implement step 1 above in polynomial time. This is done in Lemma 4.3.  $\square$

LEMMA 4.3. For a fixed  $X \in \mathcal{P}$ , one can compute  $\bar{u}_X \in \mathcal{S}^{n-1}$  such that

$$\mu(H^+(\bar{u}_X, \bar{x})) \leq \min_{u \in U_X} \mu(H^+(u, \bar{x})) + \delta \quad \text{and} \quad \|u - \bar{u}_X\|_\infty \leq \delta$$

for some  $u \in \operatorname{argmin}_{u \in U_X} \mu(H^+(u, \bar{x}))$ , using an algorithm whose running time is bounded by a polynomial in  $\log(\frac{1}{\delta})$  and the size of an irredundant description of  $P$ .

This lemma can be proved using methods from real algebraic geometry for quantifier elimination in systems of polynomial inequalities.

*Proof.* For a fixed partition  $X \in \mathcal{P}$  the feasible region  $U_X$  is described by a system of linear inequalities  $Au \leq b$  and a single quadratic equality constraint  $u_1^2 + u_2^2 + \cdots + u_n^2 = 1$ . We claim the objective function can be written as the ratio of two polynomials in  $u_1, \dots, u_n$ . Subject to these constraints, we need to minimize  $\mu(H^+(u, \bar{x}))$ . Since  $X$  is the partition of the vertices of  $P$  induced by the hyperplane  $u \cdot x = u \cdot \bar{x}$  (since  $u \in U_X$ ), the set of edges intersected by this hyperplane is fixed. Moreover, the coordinates of the point of intersection of any such edge and this hyperplane can be expressed by a ratio of linear functions of  $u_1, \dots, u_n$ . Indeed, suppose the edge intersected is the convex hull of the vertices  $v_1, v_2 \in \mathbb{R}^n$ . Then there exists  $\lambda \in [0, 1]$  such that  $u \cdot (\lambda v_1 + (1 - \lambda)v_2) = u \cdot \bar{x}$ . Thus,  $\lambda = \frac{u \cdot (\bar{x} - v_2)}{u \cdot (v_2 - v_1)}$ , and the point of intersection is  $\lambda v_1 + (1 - \lambda)v_2$  which is a ratio of linear functions of  $u$ . Also,  $P \cap H^+(u, \bar{x})$  can be decomposed into a simplicial complex whose combinatorial structure only depends on  $X$  and not on the actual values of  $u \in U_X$ . The volume of  $P \cap H^+(u, \bar{x})$  can be written as the sum of the volumes of these simplices. Since the volume of a simplex can be written as a polynomial in the coordinates of its vertices, we obtain that  $\mu(H^+(u, \bar{x}))$  is the sum of ratios of polynomials in  $u_1, \dots, u_n$  with degree bounded by a function of  $n$  only, which can be written as a single ratio of polynomials in  $u_1, \dots, u_n$ , where the degrees of the polynomials are bounded by a function of  $n$  only. Thus, finding  $u_X \in \operatorname{argmin}_{u \in U_X} \mu(H^+(u, \bar{x}))$  is equivalent to solving a mathematical optimization problem of the following type:

$$\min_{u_1, \dots, u_n} \frac{p(u)}{q(u)} \quad \text{s.t.} \quad Au \leq b, \quad u_1^2 + u_2^2 + \cdots + u_n^2 = 1.$$

The above is equivalent to the following polynomial optimization problem:

$$\min_{z, u_1, \dots, u_n} z \quad \text{s.t.} \quad p(u) = z \cdot q(u), \quad Au \leq b, \quad u_1^2 + u_2^2 + \cdots + u_n^2 = 1.$$

This optimization problem can be solved to within  $\delta$  accuracy of the objective and the solution by performing a binary search on the objective value and using quantifier elimination methods for testing feasibility of polynomial systems of inequalities and equalities. For polynomial systems with a fixed number of variables this can be done in polynomial time in the size of the coefficients [5]. See also [19, Remark on p. 2].  $\square$

**4.1.2. Counting measure on the integer points in two-dimensional polytopes.** If we use the counting measure on the integer points in a polytope, the algorithm requires no accuracy parameter  $\epsilon$ .

THEOREM 4.4. Let  $P = \{x \in \mathbb{R}^2 : Ax \leq b\}$  be a rational polytope, where  $A \in \mathbb{Z}^{m \times 2}$  and  $b \in \mathbb{Z}^m$ , such that  $P \cap \mathbb{Z}^2 \neq \emptyset$ . Let  $\mu$  denote the uniform measure on  $P \cap \mathbb{Z}^2$ . Then in polynomial time in the input size of  $A$  and  $b$ , one can compute a point

$$z \in \mathcal{C}(\mathbb{Z}^2, \mu).$$

*Proof.* By utilizing the fact that  $f_\mu$  is concave (Lemma 3.1) and Theorem 4.1, it suffices to show that for a given  $\bar{x} \in \mathbb{Z}^2$  one can compute in polynomial time

$$\bar{u} \in \operatorname{argmin}_{u \in \mathcal{S}^1} \mu(H^+(u, \bar{x})).$$

Let  $g : [0, 2\pi) \rightarrow [0, 1]$  be defined as  $g(\alpha) := \mu(H^+(\sin(\alpha), \cos(\alpha))^\top, \bar{x})$ . The key observations are that  $g$  is piecewise constant and that the domain  $[0, 2\pi)$  can be partitioned into a polynomial number of intervals  $S_i$  such that  $g$  is monotone on each of them. This implies, that in order to compute  $\bar{u}$ , one only needs to evaluate  $g$  at the beginning and the end of each interval  $S_i$ .

Let  $l^+(\alpha)$  denote the line segment  $P \cap \{\bar{x} + \lambda(\sin(\alpha + \pi/2), \cos(\alpha + \pi/2))^\top : \lambda \geq 0\}$  and let  $l^-(\alpha)$  denote  $P \cap \{\bar{x} + \lambda(\sin(\alpha - \pi/2), \cos(\alpha - \pi/2))^\top : \lambda \geq 0\}$ . Observe that  $g(\alpha)$  is monotone increasing if the line segment  $l^+(\alpha)$  is longer than the line segment  $l^-(\alpha)$  and  $g(\alpha)$  is monotone decreasing if the line segment  $l^+(\alpha)$  is shorter than the line segment  $l^-(\alpha)$ . Hence, the monotonicity can only change when the two lengths are equal. All those critical  $\alpha$  can be computed by comparing each pair of facets.  $\square$

## 4.2. Approximation algorithms.

### 4.2.1. A Lenstra-type algorithm to compute approximate centerpoints.

As we already pointed out in section 2, centerpoints can be used to design “optimal” oracle-based algorithms for convex mixed-integer optimization problems. In turn, it is possible to employ linear mixed-integer optimization techniques to compute approximate centerpoints. However, this comes with a significant loss in the approximation guarantee. Recall the definition of  $\mu_{mixed,P}$  from (7).

**THEOREM 4.5.** *Let  $n, d \in \mathbb{N}$  be fixed and let  $P$  be a rational polytope. Then in polynomial time in the input size of  $P$ , one can find a point*

$$z \in \left\{ x \in \mathbb{Z}^n \times \mathbb{R}^d : f_{\mu_{mixed,P}}(x) \geq \frac{1}{2^{n^2(d+1)(n+1)}} \right\}.$$

*Proof.* By Theorem 4.2, the statement holds for  $n = 0$ . Also, since Theorem 3.12 is constructive, there exists an  $\bar{\omega}$  that only depends on  $n$  and  $d$ , such that the theorem holds true provided that the lattice width of  $P$  is larger than  $\bar{\omega}$ .

By induction we assume that the result is true for  $n - 1$ . Further, we may assume that the lattice width is smaller than  $\bar{\omega}$ . Without loss of generality, we assume that the flatness direction of  $P$  is equal to the  $n$ th unit vector, i.e.,  $\min_{x \in P} x_n \geq 0$  and  $\max_{x \in P} x_n \leq \bar{\omega}$ . We define  $P_i := P \cap \{x \in \mathbb{R}^{n+d} : x_n = i\}$  and the corresponding uniform measure  $\mu_i$ . By the induction hypothesis, we can compute  $z_i \in \{x \in \mathbb{Z}^n \times \mathbb{R}^d : f_{\mu_i}(x) \geq \frac{1}{2^{(n-1)^2(d+1)(n)}}\}$  for  $i = 0, \dots, \bar{\omega}$ .

We define the finite auxiliary measure

$$\bar{\mu}(x) := \begin{cases} \mu(P_i) & \text{if } x = z_i, \\ 0 & \text{otherwise.} \end{cases}$$

Then, with a brute force approach, we compute the centerpoint  $z$  in  $\mathcal{C}(\mathbb{Z}^n \times \mathbb{R}^d, \bar{\mu})$ .

It remains to show that  $z \in \{x \in \mathbb{Z}^n \times \mathbb{R}^d : f_{\mu_{mixed,P}}(x) \geq \frac{1}{2^{n^2(d+1)(n+1)}}\}$ . Let  $H^+$  be any half-space containing  $z$ . Note that, for all  $i$  we have  $\mu(P_i \cap H^+) \geq$

$\frac{1}{2^{(n-1)^2}(d+1)^n} \bar{\mu}(\{z_i\} \cap H^+)$ . Hence,

$$\mu(P \cap H^+) = \sum_{i=0}^{\bar{\omega}} \mu(P_i \cap H^+) \geq \frac{1}{2^{(n-1)^2}(d+1)^n} \sum_{i=0}^{\bar{\omega}} \bar{\mu}(\{z_i\} \cap H^+) \geq \frac{1}{2^{n^2}(d+1)^{n+1}},$$

where the last inequality comes from Theorem 3.3.  $\square$

**4.2.2. Computing approximate centerpoints with a Monte Carlo algorithm.** In this section, we compute  $\epsilon$ -centerpoints, but for any family of measures from which one can sample uniformly. However, now the algorithm's runtime depends polynomially on  $\frac{1}{\epsilon}$ , as opposed to  $\log(\frac{1}{\epsilon})$  as for the uniform measure on rational polytopes from section 4.1.

Suppose we have access to two black-box algorithms:

1. OPT is an algorithm which works for some family  $\mathcal{S}$  of closed subsets of  $\mathbb{R}^n$ . OPT takes as input a closed set  $S \in \mathcal{S}$  and (approximately) computes  $\arg\max_{x \in S} g(x)$  for any quasi-concave function  $g$ , given an (approximate) evaluation oracle for  $g$  and an (approximate) separation oracle for the sets  $\{x : g(x) \geq \alpha\}_{\alpha \in \mathbb{R}}$ . Let  $T_1(S)$  be the number of calls that OPT makes to the evaluation and separation oracles, and  $T_2(S)$  be the number of elementary arithmetic operations OPT makes during its execution.
2. SAMPLE is an algorithm which works for some family of probability measures  $\Gamma$ . SAMPLE takes as input a measure  $\mu \in \Gamma$  and produces a sample point  $x \in \mathbb{R}^n$  from the measure  $\mu$ . Let  $T(\mu)$  be the running time for SAMPLE.

We now show that with access to the above two algorithms, one can compute an  $\epsilon$ -centerpoint for  $(S, \mu) \in \mathcal{S} \times \Gamma$ .

**THEOREM 4.6.** *Let  $\mathcal{S}$  be a family of closed subsets of  $\mathbb{R}^n$  equipped with an algorithm OPT as described above, and let  $\Gamma$  be a family of measures on  $\mathbb{R}^n$  equipped with an algorithm SAMPLE as described above.*

*There exists a Monte Carlo algorithm which takes as input  $(S, \mu) \in \mathcal{S} \times \Gamma$ , real numbers  $\epsilon, \delta > 0$  and computes an  $\epsilon$ -approximate centerpoint for  $S, \mu$  with probability at least  $1 - \delta$ . The running time of this algorithm is  $T_1(S) \cdot N^n + T_2(S) + T(\mu) \cdot N$ , where  $N = O(\frac{1}{\epsilon^2}((n+1) + \log \frac{1}{\delta}))$ .*

To prove this theorem, we will need a deep result from probability theory that has resulted after a long line of research sparked by the seminal ideas of Vapnik and Chervonenkis [44], and culminated in a result of Talagrand [41]. The following theorem is a rewording of Talagrand's result [41], specialized for function classes with bounded VC-dimension.

**THEOREM 4.7.** *Let  $(X, \mu)$  be a probability space. Let  $\mathcal{F}$  be a family of functions mapping  $X$  to  $\{0, 1\}$  and let  $\nu$  be the VC dimension of the family  $\mathcal{F}$ . There exists a universal constant  $C$ , such that for any  $\epsilon, \delta > 0$ , if  $M$  is a sample of size  $C \cdot \frac{1}{\epsilon^2}(\nu + \log \frac{1}{\delta})$  drawn independently from  $X$  according to  $\mu$ , then with probability at least  $1 - \delta$ , for every function  $f \in \mathcal{F}$ ,  $|\frac{|\{x \in M : f(x)=1\}|}{|M|} - \mu(\{x \in X : f(x)=1\})| \leq \epsilon$ .*

*Proof of Theorem 4.6.* We call SAMPLE to create a sample  $M$  of size  $C \cdot \frac{1}{\epsilon^2}((n+1) + \log \frac{1}{\delta})$  by drawing independently and uniformly at random from  $S$  (note that  $M$  may contain multiple copies of the same point from  $S$ ). Since the VC dimension of the family of half-spaces in  $\mathbb{R}^n$  is  $n+1$ , we know from Theorem 4.7 that with probability at least  $1 - \delta$ , for every half-space  $H^+$ ,  $|\frac{|H^+ \cap M|}{|M|} - \mu(H^+)| \leq \epsilon$ . Let  $\mu'$  be the counting measure on  $M$ . Then we obtain that  $|f_{\mu'}(x) - f_{\mu}(x)| \leq \epsilon$  for all

$x \in \mathbb{R}^n$ . Therefore, any  $x^* \in \arg \max_{x \in S} f_{\mu'}(x)$  is an  $\epsilon$ -centerpoint for  $S$ . This can be achieved by calling OPT to compute  $x^* \in \arg \max_{x \in S} f_{\mu'}(x)$ . For this, we need to exhibit evaluation and separation oracles for  $f_{\mu'}$ . But notice that, by Lemma 3.1, this can be accomplished by simply implementing the following procedure: Given  $x \in \mathbb{R}^d$ , find the best hyperplane  $H$  through  $x$  such that  $\frac{|H^+ \cap M|}{|M|}$  is minimized. This can be done in time  $O(|M|^n)$  by simply enumerating all hyperplanes that contain  $n-1$  affinely independent points from  $M$ .  $\square$

The following result is a consequence.

**THEOREM 4.8.** *Let  $n$  and  $d$  be fixed integers. There exists a Monte Carlo algorithm which takes as input an integer  $m \geq 1$ , a matrix  $A \in \mathbb{R}^{m \times (n+d)}$ , a vector  $b \in \mathbb{R}^m$ , real numbers  $\epsilon, \delta > 0$  and returns an  $\epsilon$ -approximate centerpoint when  $S = \mathbb{Z}^n \times \mathbb{R}^d$  and  $\mu$  is the uniform measure on  $\{x \in \mathbb{Z}^n \times \mathbb{R}^d : Ax \leq b\}$ , with probability  $1 - \delta$ . The running time of the algorithm is a polynomial in  $m, \log(\max\{|A_{i,j}|, |b_k|\}), \frac{1}{\epsilon}, \log \frac{1}{\delta}$ .*

*Proof.* By using classical results on maximizing quasi-concave functions over the integer points in a polyhedron [20], OPT can be implemented for the family  $\mathcal{S}$  which is the collection of all sets  $S$  that can be represented as the set of mixed-integer points in a rational polytope. For  $n = 0$ , SAMPLE can be implemented for the uniform measure on polytopes using well-studied techniques; e.g., see Vempala's survey [45]. For  $n \geq 1$ , SAMPLE can be implemented for the uniform measure on mixed-integer points in a polytope by adapting a result of Pak [39] on  $d = 0$  to  $d \geq 1$  and using results on computing mixed-integer volumes in polynomial time for fixed dimensions [3].  $\square$

**Acknowledgments.** We thank two anonymous referees whose pointers to the literature and numerous suggestions helped to improve the content and presentation of this paper. Earlier versions of this paper from [arxiv.org](http://arxiv.org) and IPCO 2016 contained results about the uniqueness of the centerpoint. We were made aware by one of the anonymous referees that such uniqueness results already existed in the convex geometry literature [46, Proposition 1]. While our uniqueness result had been obtained independently and without knowledge of these prior results, the proof ideas used by us were quite similar to the existing proof. Consequently, we do not find any value in including these results in this version of the paper.

## REFERENCES

- [1] G. AVERKOV, *On maximal  $S$ -free sets and the Helly number for the family of  $S$ -convex sets*, SIAM J. Discrete Math., 27 (2013), pp. 1610–1624.
- [2] G. AVERKOV AND R. WEISMANTEL, *Transversal numbers over subsets of linear spaces*, Adv. Geom., 12 (2012), pp. 19–28, <https://doi.org/10.1515/advgeom.2011.028>.
- [3] V. BALDONI, N. BERLINE, M. KOEPPE, AND M. VERGNE, *Intermediate sums on polyhedra: Computation and real ehrhart theory.*, Mathematika, 59 (2013), pp. 1–22.
- [4] A. BASU AND T. OERTEL, *Centerpoints: A link between optimization and convex geometry*, in International Conference on Integer Programming and Combinatorial Optimization, Springer, 2016, pp. 14–25.
- [5] S. BASU, R. POLLACK, AND M.-F. ROY, *On the combinatorial and algebraic complexity of quantifier elimination*, J. ACM (JACM), 43 (1996), pp. 1002–1045.
- [6] A. BELLONI AND R. M. FREUND, *On the symmetry function of a convex set*, Math. Program., 111 (2008), pp. 57–93.
- [7] D. BERTSIMAS AND S. VEMPALA, *Solving convex programs by random walks*, J. ACM, 51 (2004), pp. 540–556, <https://doi.org/10.1145/1008731.1008733>.
- [8] P. BRASS, L. HEINRICH-LITAN, AND P. MORIN, *Computing the center of area of a convex polygon*, Internat. J. Comput. Geom. Appl., 13 (2003), pp. 439–445.
- [9] D. BREMNER, D. CHEN, J. IACONO, S. LANGERMAN, AND P. MORIN, *Output-sensitive algorithms for Tukey depth and related problems*, Stat. Comput., 18 (2008), pp. 259–266.

- [10] A. CAPLIN AND B. NALEBUFF, *On 64%-majority rule*, *Econometrica*, 56 (1988), pp. 787–814.
- [11] T. M. CHAN, *An optimal randomized algorithm for maximum Tukey depth*, in *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, Philadelphia, 2004, pp. 430–436.
- [12] M. CONFORTI, G. CORNUÉJOLS, A. DANILIDIS, C. LEMARÉCHAL, AND J. MALICK, *Cut-generating functions and  $s$ -free sets*, *Math. Oper. Res.*, 40 (2014), pp. 276–391.
- [13] D. DADUSH, *Integer Programming, Lattice Algorithms, and Deterministic Volume Estimation*, Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA, 2012.
- [14] J. DE LOERA, R. LA HAYE, D. OLIVEROS, AND E. ROLDÁN-PENSADO, *Beyond chance-constrained convex mixed-integer optimization: A generalized calafiore-campi algorithm and the notion of  $s$ -optimization*, *J. Convex Anal.*, to appear.
- [15] J. DE LOERA, R. LA HAYE, D. OLIVEROS, AND E. ROLDÁN-PENSADO, *Helly Numbers of Algebraic Subsets of  $\mathbb{R}^d$* , preprint, arXiv:1504.00076v2, 2015.
- [16] A. DEL PIA AND J. POSKIN, *On the mixed binary representability of ellipsoidal regions*, in *International Conference on Integer Programming and Combinatorial Optimization*, Springer, 2016, pp. 214–225.
- [17] D. L. DONOHO AND M. GASKO, *Breakdown properties of location estimates based on halfspace depth and projected outlyingness*, *Ann. Statist.*, 20 (1992), pp. 1803–1827.
- [18] R. DYCKERHOFF AND P. MOZHAROVSKIY, *Exact computation of the halfspace depth*, *Comput. Statist. Data Anal.*, 98 (2016), pp. 19–30.
- [19] D. Y. GRIGOR'EV AND N. VOROBYOV, *Solving systems of polynomial inequalities in subexponential time*, *J. Symbolic Comput.*, 5 (1988), pp. 37–64.
- [20] M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *Geometric Algorithms and Combinatorial Optimization*, *Algorithms Combin.* 2, Springer, Berlin, 1988.
- [21] P. GRUBER, *Convex and Discrete Geometry*, *Grundlehren Math. Wiss.* 336, Springer, Berlin, 2007.
- [22] B. GRÜNBAUM, *Partitions of mass-distributions and of convex bodies by hyperplanes*, *Pacific J. Math.*, 10 (1960), pp. 1257–1261.
- [23] B. GRUNBAUM, *Measures of symmetry for convex sets*, in *Convexity: Proceedings Sympos. Pure Math.* 7, AMS, Providence, RI, 1963, pp. 233–270.
- [24] S. HEINZ, *Complexity of integer quasiconvex polynomial optimization*, *J. Complexity*, 21 (2005), pp. 543–556, <https://doi.org/10.1016/j.jco.2005.04.004>.
- [25] R. HILDEBRAND AND M. KÖPPE, *A new Lenstra-type algorithm for quasiconvex polynomial integer minimization with complexity  $2o(n \log n)$* , *Discrete Optim.*, 10 (2013), pp. 69–84.
- [26] A. HOFFMAN, *Binding constraints and Helly numbers*, *Ann. New York Acad. Sci.*, 319 (1979), pp. 284–288.
- [27] R. KANNAN, *Minkowski's convex body theorem and integer programming*, *Math. Oper. Res.*, 12 (1987), pp. 415–440, <https://doi.org/10.1287/moor.12.3.415>.
- [28] A. KORKINE AND G. ZOLOTAREFF, *Sur les formes quadratiques*, *Math. Ann.*, 6 (1873), pp. 366–389, <https://doi.org/10.1007/BF01442795>.
- [29] G. A. KOSHEVOY, *The Tukey depth characterizes the atomic measure*, *J. Multivariate Anal.*, 83 (2002), pp. 360–364.
- [30] J. LAGARIAS, H. LENSTRA, JR., AND C.-P. SCHNORR, *Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice*, *Combinatorica*, 10 (1990), pp. 333–348, <https://doi.org/10.1007/BF02128669>.
- [31] H. LENSTRA, JR., *Integer programming with a fixed number of variables*, *Math. Oper. Res.*, 8 (1983), pp. 538–548, <https://doi.org/10.1287/moor.8.4.538>.
- [32] R. Y. LIU, R. J. SERFLING, AND D. L. SOUVAIN, *Data Depth: Robust Multivariate Analysis, Computational Geometry, and Applications*, *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* 72, AMS, Providence, RI, 2006.
- [33] D. MICCIANCIO AND P. VOULGARIS, *A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations [extended abstract]*, in *STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*, ACM, New York, 2010, pp. 351–358.
- [34] K. MOSLER, *Depth statistics*, in *Robustness and Complex Data Structures*, Springer, Berlin, 2013, pp. 17–34.
- [35] A. NEMIROVSKI AND D. YUDIN, *Problem Complexity and Method Efficiency in Optimization*, Wiley, Chichester, England, 1983.
- [36] Y. NESTEROV, *Introductory Lectures on Convex Optimization*, *Appl. Optim.* 87, Kluwer Academic, Boston, 2004.
- [37] T. OERTEL, *Integer Convex Minimization in Low Dimensions*, Ph.D. thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich, Zürich, 2014.



- [38] T. OERTEL, C. WAGNER, AND R. WEISMANTEL, *Integer convex minimization by mixed integer linear optimization*, Oper. Res. Lett., 42 (2014), pp. 424–428, <https://doi.org/10.1016/j.orl.2014.07.005>.
- [39] I. PAK, *On sampling integer points in polyhedra*, in Foundations of Computational Mathematics, World Scientific, River Edge, NJ, 2000, pp. 319–324.
- [40] P. J. ROUSSEEUW AND I. RUTS, *The depth function of a population distribution*, Metrika, 49 (1999), pp. 213–244.
- [41] M. TALAGRAND, *Sharper bounds for Gaussian and empirical processes*, Ann. Probab., (1994), pp. 28–76.
- [42] G. TOTH, *Measures of Symmetry for Convex Sets and Stability*, Springer, Cham, Switzerland, 2015.
- [43] J. W. TUKEY, *Mathematics and the picturing of data*, in Proceedings of the International Congress of Mathematicians, Vol. 2, Canadian Mathematical Congress, Vancouver, Canada, 1975, pp. 523–531.
- [44] V. N. VAPNIK AND A. Y. CHERVONENKIS, *On the uniform convergence of relative frequencies of events to their probabilities*, Theory Probab. Appl., 16 (1971), pp. 264–280.
- [45] S. S. VEMPALA, *Recent progress and open problems in algorithmic convex geometry*, in LIPIcs-Leibniz International Proceedings in Informatics, Vol. 8, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Wadern, Germany, 2010, pp. 42–64.
- [46] E. WERNER, *Floating bodies and illumination bodies*, in Proceedings of the International Conference Integral Geometry and Convexity: Wuhan, China, 18–23 October 2004, World Scientific, Hackensack, NJ, 2006, pp. 129–140.
- [47] S. YILDIZ AND G. CORNUÉJOLS, *Cut-generating functions for integer variables*, Math. Oper. Res., 41 (2016), pp. 1381–1403.